Thesis Proper

An exploration of pragmatics in autistic spectrum disorder and normal adults and children

Table of Contents

List of Figures & Tables

Acknowledgements

For M.E.

I would like to thank Professor Dorothy Bishop for her supervision and support; Dr David Cram for discussion of the linguistic examples cited in Chapter 2; Professor Ruth Kempson for discussion of Dynamic Syntax and Theory of Mind; and Professor Robyn Carston and Dr Nausicaa Pouscoulous for discussion of experimental tests of Relevance Theory and of *ad hoc* concept construction.

Abstract: An exploration of pragmatics in autistic spectrum disorder and normal adults.

This thesis investigates pragmatic abilities in the language of people with autistic spectrum disorders ('Autistic Disorder' as per APA, 2010) and of typically developing children. The thesis introduces a framework describing the levels of pragmatic abilities, and uses this framework to discuss the existing literature, before using it to discuss a corpus of clinical-pragmatic language. The literature review and naturalistic study reveal mixed evidence for pragmatic impairment, in particular revealing an absence of evidence for impairment to high-level pragmatics inferences known as particularised conversational implicature. The thesis then reports a set of experimental studies, using a method adapted from Surian, Baron-Cohen and Van der Lely 1996), with adults with autistic spectrum disorders, typical adults, and typically developing children. These studies showed a significant difference between groups, suggesting an autistic deficit in implicature that is compatible with the notion of a developmental delay. These results support existing studies that posit a pragmatic deficit in autistic spectrum disorders but conflict with existing research on the scalar implicature in autism; the reasons for the differences in results are discussed. The results also support existing studies that show a slow developmental trajectory for implicature in typically developing children; possible methodological reasons for this apparent slow trajectory (i.e. task demands) are discussed. Although the studies tested both IQ and cognitive models of utterance processing as predictors of performance, these explanations only proved useful with typical adults. The thesis discusses alternative, linguistic theories that may be useful in future experimentation: in particular, linguistic-pragmatic Relevance Theory (Sperber & Wilson, 1986/1995).

Chapter 1: General Introduction

Introduction

This thesis aims to consider the nature of pragmatic dysfunction in autism. In considering this question, it draws on naturalistic data from children with Autistic Disorder (AD) and on experimental data from adults with AD and control groups of typical adults, 7-year-olds and 9-year-olds. This chapter introduces a definition of pragmatics, and uses it to explore the existing literature on pragmatic dysfunction in autism. It also establishes a working definition of AD in light of the American Psychiatric Association's (henceforth 'APA') proposed revisions to the Diagnostic and Statistical Manual for Mental disorders (APA, 2010). Chapter 2 presses these definitions into service in a naturalistic study, discussing conversational data from children with autistic spectrum disorders. The remaining chapters will explore implicatures, higher-order pragmatics.

Theoretical Introduction: What is Pragmatics?

In understanding any utterance, there is a division of labour between compositional semantics and pragmatics. In a complex expression, some of the meaning is a function of the lexical meanings of its constituents and the way these parts combine; but some of the meaning requires a hearer to work harder. Compositional semantics readily explains some expressions: for a sentence such as, 'The cat sat down on the mat' the meaning can largely be read off. But as soon as that sentence is uttered in a context, a hearer must ask herself, 'Which cat?' and 'Which mat?' - she must, that is, assign reference. A very similar sentence illustrates further the limits of compositional semantics: 'The cat sat on the mat'. Does the sentence describe the perfective event of a cat sitting down on the mat or the imperfective event of the cat

sitting on the mat for an extended period? Even such a simple sentence requires the hearer to work harder to recover the speaker's intended meaning; the hearer, in this case, must disambiguate the different (but, here, related) meanings of the word 'sat'. This ability to interpret a speaker's implicit intended meaning is pragmatics.

We should not underestimate the extent to which hearers need pragmatics, from the lexical level up. Researchers have long acknowledged that pragmatics is essential in assigning a value to indexical and otherwise context-sensitive expressions and to ambiguous expressions[1] (e.g. Grice, 1975). Many researchers now argue further: that compositional semantics underdetermines meaning to a striking extent. Murphy (1997; cited in Wilson and Carston, 2007) used the adjective 'fresh' to illustrate the context sensitivity of commonplace, and apparently clear, adjectives. He asked participants to name the antonyms of 'fresh' in various different contexts. His participants responded according to the context, giving such antonyms as the following:

1(a)

fresh	ANTONYMS
shirt	dirty
vegetables	rotten
fish	frozen
sheets	recently slept in
water	dirty/salt
bread	stale
air	polluted
outlook	tired
assistant	experienced
idea	old

[1] The Stanford Encyclopaedia of Philosophy distinguishes between simple indexical expressions, such as 'I', and context-sensitive expressions, which are expressions that contain an indexical expression, such as 'The man I love' (Braun, 2001). This distinction is assumed in the current thesis. Note also that indexical and context-sensitive expressions have fixed meanings but variable reference; ambiguous expressions have multiple fixed meanings and references.

Perhaps there is a core meaning of 'fresh' that is common to all of the uses listed above, although such a meaning is hard to express verbally; but in any case the meaning of 'fresh' clearly varies considerably. If the word 'fresh' is typical of language more generally – and research into so-called *ad hoc* concepts (e.g. Wilson & Carston, *ibid*) finds numerous similar cases – then word meanings vary considerably according to context. Even if we argued that contextual variation showed that words are massively ambiguous, we would have to conclude that to access the precise, contextual meaning, a hearer must use pragmatics. But we could reasonably go further: it would surely be unreasonable to posit massive ambiguity in the lexicon, because the attendant processing demands would be prohibitively costly; it would surely be more reasonable, because more economical, to posit an underspecified semantics and a complementary pragmatics that fine-tunes meanings according to context.

Just as a good case can be made for a pragmatics that operates on the meanings of the words that are actually uttered, so too can a good case be made for a pragmatics that enriches the uttered words. There are times when the verbatim meaning of an utterance seems to underdetermine what we would intuit as being said with that utterance: when, that is, the uttered words express only a minimal proposition (Recanati, 2002) to be fleshed out with additional information by the hearer to express the full proposition - what is said. Sometimes, that is, a hearer must recover an 'unarticulated constituent' of a proposition, either by saturation or free enrichment (Recanati, *ibid*). If Cherie says to John, 'That book is Tony's' then John must fill in, or saturate, the relation that holds between Tony and the book: Did Tony write the book, or did he just buy it? If Gordon says to Sarah, 'David went to

the edge of the cliff and jumped' then Sarah must fill in, or freely enrich, whether David jumped on the spot, over some obstacle, or off the cliff. Pragmatically important unarticulated constituents[2] cannot simply be read off from the world. To borrow an example of Recanati's (*ibid*), if someone says, 'Mary has been dancing' then it follows that Mary has been dancing somewhere; all events, after all, happen somewhere. But nothing turns on the location of the dancing: it is not part of what is said. Consider, in contrast, Margaret and her breakfast in the sentence, 'Margaret has had breakfast': Margaret's breakfast had to have happened sometime, but the temporal location did not have to be part of what is said; and yet, it is, for the sentence does not ordinarily mean that Margaret has had breakfast at some non-specific point in her life.

A pragmatics that operates on, and enriches, word meanings is a pragmatics that contributes to what is said with an utterance: in other words, to the literal meaning of an utterance. This point is striking. But pragmatics also concerns (and, indeed, does so more canonically) what is implied with an utterance. Consider the following example:

> 1(b)
> Context: It is early in the morning. Edward is very tired and must soon leave for a busy day at work. The following dialogue occurs:
>
> Matthew: Would you like a cup of tea?
> Edward: Tea would wake me up.

Although Edward literally says, 'Tea would wake me up' he intends to imply, 'Yes, I would like a cup of tea'. What Edward's utterance is really about is this deliberate, implicit proposition - this implicatum (Grice, 1975). We arrive, therefore, at the

[2] Here I adopt Recanati's (2002) distinction between metaphysical and communicational unarticulated constituents

following picture of pragmatics: pragmatics is the interpretation of implicit, intended meanings that are part of what is said with an utterance as well as what is implied by it.

It is only a short step from these dwellings on the nature of pragmatics to a clinical application, that is, to a definition of pragmatic dysfunction. If we take 'pragmatic dysfunction' to mean that pragmatics is impaired as a whole, then we should predict the following:

(1) Difficulties in assigning semantic values to indexical and otherwise context-sensitive expressions, and to ambiguous expressions: for example, deictic expressions such as pronouns or tensed expressions, and lexical ambiguity.
(2) Difficulties in understanding contextually calibrated lexical items: for example, *ad hoc* concepts such as context-sensitive uses of adjectives (e.g. 'fresh').
(3) Difficulties in saturating and freely enriching minimal propositions
(4) Difficulties in understanding implicatures[3].

We can reasonably imagine that all of predictions (1) to (4) might hold, but also that only some of (1) to (4) might hold; all of pragmatics need not be supported by exactly the same mechanisms. Accordingly, some theorists suggest that pragmatic abilities diverge: Recanati (2003), for instance, argues that there are primary pragmatic processes, which largely handle saturation and free enrichment and are non-conscious and associative, and secondary pragmatic processes, which handle other pragmatic phenomena, are conscious, and are inferential. Granted that this putative divergence is controversial, with other theorists arguing for a single

[3] For ease of exposition, I collapse metaphors and irony into implicatures. I do not mean to imply that metaphors and irony are strictly equivalent to implicatures, and hence also to each other: just that understanding a metaphorical or ironical utterance requires the hearer to recover an implicature, perhaps, for example, triggered by an ad hoc concept in the case of metaphor (Robyn Carston, *personal communication*; see Wilson & Carston, 2007) or by the echoic use of a previous proposition (Sperber & Wilson, 1986/1995).

inferential pragmatic mechanism (e.g. Sperber & Wilson, 1986/1995), we should nevertheless be alive to the possibility that predictions (1) to (4) might dissociate. Presumably most pragmatic theories would want, for instance, to distinguish between (1) to (4) in terms of processing difficulty: (1) being the simplest processing, (4) the hardest. Potentially, then, an answer to the question of, 'Which of (1) to (4) hold?' will cast light on the nature of autistic spectrum disorders, and suggest valuable avenues of research.

Pragmatics and Normal Development

An exhaustive survey of pragmatic development is beyond the scope of this thesis; however, a brief précis will prove useful here. Developmental Pragmatics shows a trajectory from naivety in children to sophistication in adults. Probably the best evidence for this trajectory comes from the development of scalar implicatures, which will be reviewed in some detail in Chapter 7. Suffice it to say that although some children are considered pragmatically naïve, the age at which sophistication starts to develop is disputed, with some researchers suggesting that 69% of 9-year-olds are pragmatically naïve (Noveck, 2001) and others suggesting that children as young as 5 show some competence (Papafragou & Musolino, 2003) and even as young as 4 with specific training (Papafragou & Tantalou, 2004). Supporting evidence for such a developmental trajectory comes from studies into metaphor comprehension (Noveck, Bianco & Castry, 2001; Winer et al, 2001) and from the resolution of ambiguous sentences such as 'Every horse did not jump the fence', for which adults prefer the reading 'Not every horse jumped the fence' (Musolino, Crain & Thornton, 2000; Noveck, Guleminger, Georgieff & Labruyere, 2006) and children either prefer the reading 'No horse jumped the fence' (Noveck et al, *ibid*) or show an ambiguity between readings (Musolino & Lidz, 2000). A

methodological study by Pouscoulous, Noveck, Politzer and Bastide (2007) has suggested that child competence depends on experimental factors such as a task's meta-linguistic demands and the presence of foils. Thus, although the age varies at which pragmatic sophistication can be demonstrated, meaningful comparisons can be made between groups if equivalent methods are used.

Pragmatics and Impairment

Pragmatics features in various disorders. Pragmatic dysfunction is present perhaps most famously in autistic spectrum disorders; indeed, dysfunction is, in effect, present by the clinical definition of that disorder. But, on some sense of the term[4], pragmatics is also implicated in right hemisphere brain damage (Mitchell & Crow. 2005), traumatic brain injury (Martin & McDonald, 2003) and schizophrenia (Langdon, Davies & Coltheart, 2002). Pragmatic dysfunction has also been demonstrated in some children with Williams Syndrome, a finding that somewhat contradicts the popular impression that children with Williams Syndrome are socially and communicatively gifted (Laws & Bishop, 2004). Consequently, we can reasonably view abnormal psychology and cognitive neuropsychology as fertile ground for the study of pragmatics.

Autism, Asperger Syndrome and DSM-V

This thesis uses the term Autistic Disorder (AD) to refer to those disorders commonly known as autism, Asperger Syndrome, childhood disintegrative disorder, and pervasive developmental disorder- not otherwise specified (PDNOS). It does so in line with recent proposals by the American Psychiatric Association (APA, 2010)

[4] The authors reported here use various definitions of the term 'pragmatics'. All are different from the definition outlined in this thesis; of these authors, Langdon, Davies and Coltheart's (2002) definition is the strictest and the closest to ours.

to streamline diagnostic categories of autistic spectrum disorders. However, the thesis cites various papers that distinguish between autism and Asperger syndrome in line with past versions of the Diagnostic and Statistical Manual of Mental Disorders. The latest published version of the manual (DSM-IV-TR, 2000) defines autism in terms of the so-called triad of impairments: (1) qualitative impairment in social interaction and (2) in communication, and (3) restricted repetitive patterns of behaviour, interests and activities. Moreover, there must also be delay or abnormality in social interaction, social use of language or symbolic or imaginative play with onset prior to 3 years' of age. This same version of the manual (*ibid*) distinguishes Asperger Syndrome from autism most notably in terms of a lack of clinically significant delay in language, cognition, self-help skills, adaptive behaviour and curiosity about the environment in childhood. In proposing its revisions, however, the APA (2010) noted that language skills in Asperger Syndrome are hard to estimate retrospectively, especially given that the syndrome is typically diagnosed quite late (at 7 years' of age), and that the seemingly distinctive characteristics of autism, with respect to Asperger Syndrome, are not necessary criteria for a diagnosis of autism. The APA (*ibid*) noted, further, two rival positions around which consensus is forming in the literature. On the one view, Asperger Syndrome is a form of high-functioning autism characterised by good formal language skills and verbal IQ. On the other, Asperger Syndrome is a separate disorder characterised by onset patterns and early language that are different from other autistic spectrum disorders. However, the APA (*ibid*) currently holds that there is no strong evidence supporting additional criteria to distinguish Asperger Syndrome from high-functioning autism, and accordingly recommends the subsuming of Asperger Syndrome into a category of Autistic Disorder, alongside

autism, childhood disintegrative disorder and pervasive developmental disorder-not otherwise specified (PDNOS). This streamlining of the diagnostic categories will doubtless be controversial. But bearing in mind the APA's persuasive reasoning, this thesis adopts the following terminology: the category of Autistic Disorder (AD) is assumed; where the term Autistic Spectrum Disorder (ASD) is used, its meaning is equivalent to Autistic Disorder; where the terms 'autism' and 'Asperger Syndrome' are used, their usage reflects other authors' use in line with published DSM criteria, and does not imply this author's commitment to a clinical distinction between the disorders.

Literature Review of Pragmatics and AD/ASD

There is variable coverage of pragmatic impairment in the empirical literature. This review summarises the literature using the definition of pragmatics given above.

(1) Difficulties in assigning semantic values to indexical and otherwise context-dependent expressions

The empirical literature gives the greatest coverage to this level of pragmatic impairment. This review briefly lays out evidence for autistic impairment in deixis and ambiguity resolution.

Deixis

Let a deictic expression be one which grammaticalizes features of the context of an utterance or speech event by acting as a place-holder for some entity given by the

context (Levinson, 1983:54). Canonical deictic phenomena are demonstratives, pronouns, tense, and spatial and temporal adverbs (Levinson, *ibid*).

The richest source of data on deixis in autism is in pronouns, especially personal pronouns. Pronouns can be used to refer in different ways. (1) They can be used endophorically, in which case the pronoun stands for a linguistic expression that is present in the broader discourse. An endophoric pronoun is also anaphoric if it refers to an expression that precedes it, as in *[The person]$_i$ writing this sentence is a psychologist.[He]$_i$ wishes to exemplify a point.* An endophoric pronoun is also cataphoric if it refers to an expression that follows it, as in *[It]$_i$ can be difficult [to find examples of cataphora]$_i$.* (2) Pronouns can also be used exophorically, in which case they stand for entities that are present in the real world. Pronouns used in this way are deictic, as in (without preceding discourse) *He is the President of Burundi*.

Fine et al (1994) conducted experiments on exophoric and endophoric expressions with children with higher-functioning autism, children with Asperger's Syndrome, and a control group of children with non-specific social problems. Ten-minute interviews were conducted with each child by a clinical psychologist, blind to the nature of the study, on a range of topics such as school, family, hobbies and vacations. These interviews were transcribed and coded for referential expressions by a linguist, again blind to the nature of the study. The code comprised the following categories: endophoric; exophoric where (a) the reference is first-person speech role, (b) second-person speech role, or (c) non-speech role; a cultural reference (e.g. 'the president', 'the queen'); a bridging reference (e.g. 'I saw John's new house. The door was painted green.'); and unclear reference, which can,

further, be ambiguous (there are multiple possible antecedents) or additioning (there are none). This code could arguably have been better arranged such that cultural references were a subcategory of exophoric reference, bridging reference a subcategory of endophoric reference, and unclear reference decomposed into unclear endophoric (ambiguous) and exophoric (additioning) reference. The code would also have benefitted from the addition of anaphoric and cataphoric categories for endophoric reference, since referring backwards and projecting forwards to an as yet unspoken constituent may call on distinct cognitive abilities.

Nevertheless, the study yielded intriguing results. These results were as follows. Children with higher-functioning autism made significantly fewer endophoric references than the Asperger's and control groups, which did not differ significantly from each other. The Asperger's and control groups, in turn, made significantly fewer exophoric references than the autistic group; again the Asperger's and control groups did not differ significantly from each other. The Asperger's group made significantly more unclear references than did the other two groups, and these unclear references were the additioning references, not the ambiguous references. The autistic group made significantly more cultural references than the other groups. Bridging references and exophoric speech roles did not differ significantly across groups. The authors interpreted their results as suggesting that children with autism make greater use in conversation of reference to the physical world compared with children with Asperger's and with non-specific social disorders, and that an apparent lack of cohesion in autistic discourse may be attributable to a reduced capacity for endophoric reference. This interpretation is persuasive, but conflicts somewhat with other research on pronoun usage.

Tager-Flusberg (Tager-Flusberg et al, 1990; Tager-Flusberg, 1993) has presented data that suggest that the very types of language that Fine et al's (1994) autistic group favour should prove problematic. Tager-Flusberg's children with autism reversed a tenth of all personal pronouns, a proportion that is three times higher than for typically developing children. Data were gathered from six children with autism and six with Down's syndrome as a control group. Five out of six children with autism were higher functioning. All children were video-recorded over a period of between one and two years for an hour each month during conversations with their mother. Some constancy was attempted in the sessions by having the experimenter give the child a present and having the mother and child play with this present. These conversations were analyzed linguistically for Mean Length of Utterance, Productive Syntax, lexical diversity and form class distribution. The relevant data for present purposes are a subset of the last category, showing that the children with autism, unlike their counterparts with Down's syndrome, made numerous pronoun reversal errors. The children would refer to themselves as 'you' and their mothers as 'I' (Tager-Flusberg, 1993). There thus seems to be a marked difficulty with a subset of deictic/exophoric pronouns.

It is unclear how contradictory Fine et al's (1994) and Tager-Flusberg's (Tager-Flusberg et al, 1990; Tager-Flusberg, 1993) are. For one thing, Fine et al's (1994) data do not, in fact, speak to the question of whether deictic pronouns were used correctly by the autistic group, since any errors would presumably have been coded as 'unclear references' and Fine et al (1994) do not describe the types of errors in enough detail to allow an analysis of pronoun reversal. It is also possible that Fine

et al's (1994) and Tager-Flusberg's (1993) samples differed in ability as a function of the severity of the autism. Tager-Flusberg (1993) herself suggests that her sample were lower functioning than in other studies (Baron-Cohen, 1989; Hobson, 1990; Leslie & Frith, 1988) in explaining a possible discrepancy between the findings of those studies and her own. This explanation – in its original context and in the present context – loses much of its force, however, without specific comparisons of participants' abilities. Perhaps a more convincing explanation is that Tager-Flusberg's (1993) autistic group was simply younger than Fine et al's (1988), and that, as Tager-Flusberg (*op cit*) notes, successful pronoun usage is only delayed, and not absent, in her sample. Thus, while it may seem surprising that the recently fragile and delayed ability evidenced by Tager-Flusberg (1990, 1993) becomes the privileged ability evidenced by Fine et al (1994), this makes rather better sense if viewed as the early stages of a delay in more sophisticated pronoun use.

The studies discussed here, though impressive, gloss over some subtle points and thereby undermine their conclusions. The Tager-Flusberg (1993) study provides the clearest example. Tager-Flusberg (*ibid*) suggests that pronoun reversal is a function of a deficit in conceptual role-taking: it is, on this view, the child's inability to understand the roles of him-/herself and the mother in the exchange that causes the confusion. This may be so, and it would be very interesting if it were so; but there is inadequate information to judge whether the explanation will hold. It is indeed interesting that the reversals are reversals of personal pronouns and not of deictic pronouns more generally; this lends weight to the view that the deficit implicates a failure to understand something about people. But it is a fact about the English language that the personal pronoun system ('I', 'you', 's/he', 'we', 'they') is more

complex than the pronoun system for objects or non-human creatures (arguably ‘this’ and ‘that’, and more surely ‘it’). This is a confound which will not exist in some other languages however - in German, for instance, both objects and non-human creatures are labelled with the same gendered pronouns as are people - which suggests that a cross-linguistic study would be justified involving speakers of a language with a more complex gender system for non-human animate and inanimate entities.

Even if we leave this confound aside, there is an interpretative dilemma for both the studies that were discussed above: it is unclear whether problems with pronoun usage resulted from an impaired ability for pronoun use *per se* or in concomitant abilities such as the ability (a) to perform the conceptual role-taking in discourse, (b) to remember and refer back to preceding discourse, or (c) to plan ahead and refer to forthcoming discourse. Since both studies used language produced by children with autism, and did not assess their receptive ability to understand pronouns when used by others, it is also unclear whether deficits are, again, in the ability for pronouns *per se* or in a sub-ability to use pronouns as opposed to understanding them.

Nevertheless, there is evidence that other types of deixis are problematic in autism; accordingly, we can argue that it is the deictic nature of these expressions, and not deficits in other supporting cognitive abilities, that causes problems. Only the briefest sketch of these data can be given here because of limits on space. In a larger study to partition autism into subtypes based on various cognitive and linguistic skills, Tager-Flusberg and Joseph (2003) used probes to elicit regular and irregular

verbal morphology. An impaired subtype produced significantly more bare verb stems than a normal subtype, crucially only in the past tense so that the problem does not seem to be a straightforward lack of verbal morphology. Perkins et al (2006) analyzed a naturalistic sample of language from 7 adults with autistic spectrum disorder, finding problems with tense, aspect, syntagmatic semantic relations and discourse coherence. For example, one speaker produced the sentence, 'I usually go to Claremont when I was a little girl': note the clash between the tenses of 'go' and 'was'; perhaps the habitual aspect of 'usually' short-circuits the assignment of correct tense morphology. Intriguingly the authors found problems only with verbal expressions, with nominal expressions being unimpaired; perhaps, as the authors suggested for other apparently correct temporal expressions, the speakers with autistic spectrum disorder had learnt some temporal expressions by rote, and repeated them correctly but formulaically; this would seem an easier feat with nominal expressions than verbal ones. Perkins et al (*ibid*) also reported problems with spatial expressions[5]. Finally, Baltaxe (1977) reported that a sample of German children with autism did not grasp the correct usage of the German second-person singular pronouns 'Sie' and 'du', where 'Sie' encodes respect or social distance (i.e. you would use it with non-familiars) and 'du' encodes familiarity.

Deficits in the production and comprehension of deictic pronouns do seem to offer good evidence for impairment at the first level of pragmatic abilities. Converging data is offered by deficits in disambiguation.

[5] Their coverage of this issue is elliptical however: they present minimal data on spatial terms, and the most likely coding category into which spatial errors would fall is prepositional errors. We would need considerably more information on the prepositional errors to judge whether they imply problems with spatial deixis, since prepositions need not imply spatial reference (e.g. 'I am thinking about prepositions')

Disambiguation

A word (or phrase) is ambiguous if it has more than one sense; strictly, if the multiple senses are unrelated then a word is ambiguous, and if they are related then a word is polysemous, but both categories will be collapsed here into the term 'ambiguous'. To exemplify the importance of disambiguation, consider the following senses of the function word 'the': a referential sense as in, 'The man standing by the fire-place is very tall'; an attributive sense as in, 'The murderer of JFK must have been a very good shot'; and a generic sense as in, 'The soldier is generally a noble type of man'. Given that we find ambiguity in the definite article, which belongs to a class of words (i.e. function words) that are generally thought to be 'meaning-light', we are led to expect even greater ambiguity with words whose lexical entries are dense with possible meanings.

We expect a deficit in disambiguation when we read, as in Weak Central Coherence Theory, that people with autism struggle to make use of the linguistic context: see, for example, deficits in selecting the contextually appropriate pronunciation of homographs[6] (Frith & Snowling, 1983; Happé, 1997). There is, indeed, ample evidence for a deficit in disambiguation. Joliffe and Baron-Cohen (1999), for instance, investigated, amongst other skills, lexical and syntactic disambiguation. They presented sentence pairs: one sentence was a context, which primed either a common or a rare interpretation of the ambiguity (the frequency being previously determined by 100 normal native speakers); the other was an ambiguous sentence,

[6] Though see Hala, Pexman and Glenwright (2007), who, on the strength of a task which presents ambiguous words twice with different senses, characterize the deficit instead as one of perseveration of the first recalled sense, hence one of executive function not central coherence.

whose ambiguity was either lexical or syntactic. Participants were asked to choose one of three possible responses to a question: one answer was contextually appropriate; one was contextually inappropriate but a possible reading of the ambiguity; the last was erroneous. As a control task, participants had to repeat pairs of sentences using the rare interpretations. Joliffe and Baron-Cohen (*ibid*) found that participants with autism made significantly more errors than participants with Asperger's and typical participants, which groups were not significantly different although there was a non-significant trend for participants with Asperger's to make more errors. There was no difference in error rates between lexical and syntactic ambiguities.

Norbury (2005) offered finer-grained supporting evidence that disambiguation is sometimes problematic in autism. She studied 9-17 year-olds with language impairment (LI), with autism and language impairment (ALI), with just autistic spectrum disorder (ASD), and typical controls. Participants with LI were recruited from specialist schools; their language impairment was assessed with the British Picture Vocabulary Scales, The Concepts and Directions subtest of the Clinical Evaluation of Language Fundaments, and the Recalling Sentences subtests of the same. Autistic features were assessed by having parents complete the Social Communication Questionnaire. Norbury used a test case of ambiguous words with dominant and subordinate meanings. A first task explored participants' knowledge of dominant and subordinate meanings; LI and ALI participants knew fewer subordinate meanings. A second task explored whether contexts appropriately facilitated the intended reading of a word and whether alternative meanings were appropriately suppressed. For example:

1(c)	*Facilitation item*	Bill ran from the bank (neutral sentence) Bill stole from the bank (biased, dominant) Bill fished from the bank (biased, subordinate)
1(d)	*Suppression item*	Bill stole from the bank (dominant) Bill fished from the bank (subordinate)

After hearing each sentence, participants saw a picture and had to decide whether the picture fit the meaning of the sentence. For facilitation items, the pictures were congruent, so that participants should answer 'yes'; for suppression items, pictures were incongruent but related to the ambiguous word, so that participants should answer 'no'. In the suppression task there were control items with unambiguous words. With facilitation items Norbury found that all participants were more accurate with dominant meanings, and that they were more accurate and faster with biased contexts. Language-impaired participants (with or without autism) were less accurate and less speedy than other participants. With suppression items Norbury found that all participants could reject incongruent pictures with unambiguous words, and that all participants were slower with ambiguous words. Language-impaired participants (with or without autism) made more errors. Norbury argued that because the two deficits – in facilitation and suppression – occurred together, language-impaired participants (LI and ALI) in the suppression items cannot build enough of a context to perform successfully, and therefore behave as though the task were a word to picture matching task. Crucially, here, it is the language impairment, and not the presence of autistic features, which seems to be to blame for deficits detected in Norbury's experimental task: as Norbury notes, this fact counts against Weak Central Coherence Theory, according to which a deficit in using context is definitional of autism. It is thus unclear whether a deficit in disambiguation, and accordingly pragmatic dysfunction, can properly be attributed to autism *per se*.

Noveck, Guelminger, Georgieff and Labruyere (2007) focused on ambiguity at a higher level. Rather than addressing lexical ambiguity, they investigated ambiguity in logical form[7], using a type of expression known as 'Every…not' sentences. Consider the ambiguous sentence, 'Every horse did not jump over the fence'. The ambiguity is neither lexical nor syntactic. The ambiguity is rather in the logical form and depends on the scope of the quantifier and of the negation. If the universal quantifier ('Every') takes scope over the negation, then the following interpretation holds: 'All horses are such that they did not jump over the fence' or, more idiomatically, 'No horse jumped over the fence' (Noveck et al, *ibid:* 74). If the negation takes scope over the universal quantifier, then the following interpretation holds: 'It is not the case that all horses jumped over the fence' (*ibid*).

Noveck et al embedded 'every…not' sentences into the following task. Participants heard stories and saw cartoons representing the events in the stories. In each story there were three possible agents or objects. After each story participants would hear a quantified statement or a factual statement, on which they performed a truth-value judgement. There were filler stories in which all agents or objects satisfied a given description. But the crucial stories were those in which two out of three agents or objects satisfied a given description. Alongside 'every…not' sentences were controls of quantified sentences without negation and (as mentioned above) factual

[7] In a sense, Noveck et al's interests are diametrically opposed to mine. They assume pragmatic dysfunction in autism, and use autism to test whether pragmatics is involved in interpreting ambiguities in 'Every…not' sentences. I make no such assumption. Indeed, I may seem to be assuming that pragmatics is essential in resolving logical-form ambiguities in 'Every…not' sentences, and using an autistic deficit in resolving such ambiguities to argue for pragmatic dysfunction. I do not intend to commit myself fully to this assumption; I simply present the deficit as possible, but equivocal, evidence of pragmatic dysfunction.

questions, to test, respectively, comprehension of quantification and comprehension of the story. For example:

1(e) All the children are not in the pool (Every…not)
1(f) All the children are in the pool (Control without negation)
1(g) The children in the pool were playing with a ball.
(Factual statement; some true, some false)

Noveck et al's participants were typical adults, typically developing children (mean age 4;7, other measures unavailable), and participants with autism (mean age 16;3, mean mental age 8;9, mean IQ 53, mean verbal IQ 58, 67%[8] passed false-belief tests). In response to such stories, participants' behaviours diverged: typical adults preferred the 'not every' reading of the sentence; typical children and participants with autism were at chance, but passed control tests. Noveck et al's data is intriguing but hard to assess for present purposes[9]. The authors give no scores on standardized tests for adults or typically developing children. Accordingly it is unclear whether the adults are typical (rather than, say, highly gifted, and atypical of population performance on this task) and indeed whether the children are genuinely typical, as reported. This omission means that it is difficult to establish the extent of any autistic deficit; rather, such a deficit is assumed by Noveck et al (2007) *en route* to testing the involvement of theory of mind. Relatedly, the fact that Noveck et al do not report, hence presumably did not do, standardized language tests makes us recall Norbury's (2005) finding that only those participants with autism who were also language-impaired showed deficits in disambiguation. We cannot tell, for sure,

[8] There is actually an ambiguity in Noveck et al's data: they report, under the label 'per cent' passing false-belief tasks the figure '.67'. If this figure implies .67% then .1 participants passed a false-belief task, which does not seem a possible outcome. If the figure implies 67%, then 10 participants passed a false-belief task, which seems more likely.

[9] Given Noveck et al's (2007) different approach, reported in footnote 6, it is unsurprising that they did not carry out extensive standardised tests. The criticisms reported here, then, are not, strictly speaking, criticisms of Noveck et al's research, but rather limitations on the application of their findings to the present debate.

whether Noveck et al's participants with autism were language impaired or not, but their low verbal IQ is suggestive. Perhaps deficits in processing of 'every…not' sentences, like lexical disambiguation, are not characteristic of autism as a whole but rather of an autistic subtype. The data are nevertheless suggestive.

To summarize the findings on ambiguity, there are grounds to suppose that disambiguation poses challenges to speakers with autism, both at the word level and at the level of logical form. It is unclear, however, whether such difficulties are endemic to the autistic population as a whole or just to those speakers who fall in the intersection of autism and language impairment. Joliffe and Baron-Cohen's (1999) and Noveck et al's (2007) data do not speak to this question. It is also worth noting that none of the studies described here demonstrate an unequivocally pragmatic impairment. While pragmatics may be crucial in disambiguation, there are doubtless also more structural elements at work; indeed, Norbury's (2005) findings in her first experiment implicate these very structural elements. We can easily conceive of a pragmatics in autism that is intact but that has different inputs and produces, accordingly, different outputs. Subtle structural differences might, therefore, give rise to apparent pragmatic dysfunction.

(2) Difficulties in understanding contextually calibrated lexical items: for example, ad hoc concepts such as context-sensitive uses of adjectives

While numerous studies have investigated disambiguation in autism, none has directly investigated whether speakers with autism can contextually calibrate lexical items. There are, however, hints of difficulties in studies treating other phenomena, such as disambiguation, concept formation and cognitive flexibility. Problems with disambiguation can potentially be reinterpreted as problems in calibrating lexical

items to a context. This reinterpretation is motivated theoretically in that, in many cases, such as the context-sensitive uses of 'fresh', there is simply little point in insisting that the many senses are all encoded. This possibility - that at least some cases of apparent problems with disambiguation are actually problems with contextual calibration - will be taken up in the next chapter, on novel corpus data. The possibility is supported by certain relevant experimental findings. Various researchers (e.g. Barsalou, 1993; Wilson & Carston, 2007) have stressed a link between flexible uses of concepts in reasoning tasks and flexible uses of words. There is ample evidence of autistic deficits in concept formation and flexibility. Minshew, Meyer and Goldstein (2002), for instance, used various reasoning tasks, such as the 20 Questions Task, to show that participants with autism had difficulties in forming concepts but did not have difficulties in identifying concepts. Bishop and Norbury (2005) showed similar findings with tests of ideational fluency and, furthermore, linked these tests to language. The authors had participants suggest as many different uses as possible for certain specified objects and as many meanings possible for arbitrary line drawings, finding that participants who scored low on ideational-fluency tests also showed communicative abnormalities. These are tantalising hints of possible problems with flexible, contextual uses of words; but so far, the experimental data are lacking to cement the hints.

(3) Difficulties in saturating and freely enriching minimal propositions

The literature on level (3), as with that on level (2), is sparse, and even that sparse literature shows mixed data. Two studies have explicitly explored autistic language

at what this thesis has called level (3) pragmatics: De Villiers, Stainton and Szatmari (2007) and De Villiers, Myers and Stainton (*forthcoming*). Both sets of authors drew on the same corpus of autistic language: a resource that was composed of transcribed semi-structured conversations with 12 speakers with autism aged 10-16 who had previously been identified when they were aged 4-6. These speakers with autism all scored well on tests of vocabulary and expressive language. The corpus was coded for grammatical errors, breakdowns in communication, and failures to take into account the researcher's needs as a hearer. The coding criteria identified pragmatic problems in fixing on the literal meaning of the utterance, and comprised the following categories: indexicals, sub-sentences, null complements, possessives, quantifier domain restriction, scalar adjectives, homophony and polysemy. The categories 'sub-sentences' and 'null complements' correspond closely to level (3) pragmatics, and so are focused on here: sub-sentences seem to require massive free enrichment; null complements seem to be more or less equivalent to saturation.

To exemplify the pragmatic abilities of some speakers with autism, witness the following extracts:

1(h) **Sub-sentences**
Researcher: what kinds of things did you miss?
Participant: **math**.
Researcher: math?
Participant: I don't like math
Researcher: mmhm
Participant: **science**
Participant: I like making experiments.
Participant: **but alone**

1(i) **Null complements**
Participant: I think I was twelve four fourteen I started **playing**.

Participant: and then **played** for about two or three years.
Participant: **practised up**.
Participant: I got **good enough**.
Participant: and then I **joined**.

(De Villiers, Stainton and Szatmari, 2007)

In these examples, the participants (with autism) seem to successfully negotiate level (3) pragmatics. The participant in example 1(h) successfully uses sub-sentences to respond to the researcher's question: each of the boldfaced utterances is a sub-sentence. The participant's success is not threatened by the researcher responding, 'Math?' as long as the researcher is understood to be requesting more information rather than clarification. The participant in example 1(i) successfully uses null complements. His utterances occur in the context of a discussion of basketball, so that he clearly means that he started playing basketball when he was twelve or fourteen, that he played basketball for two or three years, and that he practised basketball until he was good enough at basketball to join a basketball league.

To consider the pragmatic difficulties of some speakers with autism, consider the following extracts:

1(j) **Sub-sentence**
Researcher: who's in your family?
Participant: hmm I don't know.
Researcher: are there five of you?
Participant: yes.
Participant: **my cat**

1(k) **Null complement**
Researcher: Do you have a sister?
Participant: Yes and she **won**!
Researcher: What did she win?

(De Villiers, Myers and Stainton, *forthcoming*)

In these examples, the speakers seem to show dysfunction in level (3) pragmatics. The speaker in example 1(j) inappropriately uses the sub-sentence ‘my cat’, so that his meaning is unclear: interpreted charitably, his response means that his cat is a member of the family; but is his cat one of the five or an additional member? The speaker in example 1(k) inappropriately uses a null complement, so that his meaning is likewise unclear: as the researcher asks, what did his sister win?

These data are important and suggestive, but fragmentary. They suggest that some, but not all, speakers with autism struggle with level (3) pragmatics. But because the data result from a single study with few participants, they bear limited scrutiny. While the data are reminiscent of the dys-fluencies and misunderstandings shown by pragmatically dysfunctional children in Bishop and Adams (1992) and Bishop, Chan, Adams, Hartley and Weir (2000), to be discussed in the next chapter, the literature more generally does not seem to have explored level (3) pragmatics. The data are interesting enough to suggest that future studies could valuably focus on level (3) pragmatics.

(4) Difficulties in understanding implicatures

Given that there are differences between autistic and non-autistic populations in lower-level pragmatics, similar differences at the higher, and presumably more taxing, implicature level seem likely. But in the existing literature such differences are not clearly evidenced. This point must immediately be qualified, for in this thesis implicatures are taken to include metaphor and irony, and an autistic deficit in understanding these phenomena is considered some of the most striking evidence

for pragmatic dysfunction (Happé, 1993). Collapsing implicature, metaphor and irony together is principled, because implicatures do seem to deliver the meaning of metaphors and ironies. But because metaphoric and ironic implicatures are triggered by some other phenomenon, such as an ad hoc concept or an echoic use of a previous proposition (Robyn Carston, *personal communication*; see Wilson & Carston, 2007), the implicatures cannot be responsible alone for the metaphor or irony. Accordingly, problems with implicatures might account for problems with metaphors and ironies; but so too might problems with the metaphoric or ironic trigger. When implicatures are considered in isolation in autistic spectrum disorders, implicatures seem less problematic.

There is seemingly only one study on autistic spectrum disorders that has focused on the implicature. In this study, Pijnacker, Hagoort, Buitelaar, Teunisse and Geurts (2009) explored scalar implicatures in speakers with Asperger's Syndrome (AS) or high-functioning autism (HFA)[10]. In so doing they followed a trend in experimental pragmatics, where scalar quantifiers ('some' implicates 'not all'), logical operators ('or' implicates 'and not both') and modal verbs ('might' implicates 'and not must') have been used in various methodologies with adults and children. Pijnacker et al's materials are exemplified by the following items:

[10] Pijnacker et al (2009) distinguished autism from Asperger's Syndrome conceptually with the criterion of language delay, arguing that children with autism, but not Asperger's, show a clinically significant delay in the production of first words by the age of two and first phrases by the age of three (DSM-IV, 1994). The authors distinguished the disorders procedurally by, first, conducting an expert clinical evaluation using the DSM-IV criteria (DSM-IV, *ibid*) and, second, confirming the diagnosis with the autism diagnostic interview-revised (Lord et al, 1994) – a semi-structured interview with parents, including questions about the clinical participant's early behaviour.

1(l)	*True Universal*	'All sparrows are birds' *logically true*
1(m)	*Under-informative*	'Some sparrows are birds' *pragmatically false*
1(n)	*False Universal*	'All birds are sparrows' *logically false*
1(o)	*True Existential*	'Some birds are sparrows' *logically true*

As Pijnacker et al argue, a strictly logical participant should reject only the false universal, but a pragmatically able participant should also reject the under-informative item, because it implicates the false proposition, 'Some sparrows are not birds'. Pijnacker et al used quantifiers, as above, and disjunction (i.e. 'or'); they controlled all materials for length, and presented them pseudo-randomly among filler items. Participants performed truth-value judgements on the sentences.

Pijnacker et al found that, overall, the performance of controls and participants with ASD (i.e. the HFA and AS groups combined) did not differ significantly. But their autistic-spectrum groups differed significantly from each other. Their adults with AS were generally as fast and pragmatic as controls, and even faster than controls with the scalar term 'some'. Their adults with HFA were slower and more likely to accept under-informative sentences than were adults with AS. Moreover, Pijnacker et al found that, unlike the performance of the other groups, that of adults with HFA was significantly correlated with verbal intelligence: adults with HFA who scored highly on verbal intelligence produced more scalar implicatures. The authors also probed the groups' language skills with the comprehension subtest of the Aachen Aphasia Test (Graetz et al, 1992), the Offline Observation Instrument for Grammatical Comprehension (Wassenaar & Hagoort, 1994), and the Auditory Discrimination Task – subtest C (Crul & Peters, 1976). Their analyses revealed that adults with HFA scored significantly lower on both tests than the other groups,

although the scores were well above thresholds for language impairment. Language skills did not correlate with the number of logical responses.

From these data Pijnacker et al (*ibid*) drew the following guarded conclusions. Because their combined ASD (i.e. HFA and AS) sample did not differ significantly from the controls, but their AS sample differed significantly from the HFA sample, the authors suggested that the ASD population overall may not be pragmatically impaired but that AS and HFA may diverge. Noting that HFA performance was correlated with verbal intelligence, the authors suggested, however, that adults with HFA use their verbal intelligence to compensate for pragmatic difficulties, an explanation that is consistent with longer reaction times and lower pragmatic accuracy because of the implied alien mechanism. And yet, as the authors observed themselves, how verbal intelligence relates to implicature is unclear[11]; and because the authors did not distinguish between the components of verbal intelligence, any explanation will be firmly speculative. The authors also argued that theory of mind might explain their data: on their view, to produce the scalar implicature 'not all' from 'some', a hearer must assume that the speaker would have said 'all' if s/he had known that the word could be correctly applied, that is, hearers must apply first-order theory of mind. The authors drew a link between the average or above average intelligence of their participants and the common finding that people with ASD, and particularly those with AS, and average or above average evidence can pass tests of first-order, and sometimes second-order, theory of mind. Because the authors' explanation implies a systematic difference in theory of mind between their adults

[11] We could speculate that a faculty of non-demonstrative inference supports scalar implicatures and is tapped into by tests of verbal intelligence.

with HFA and their adults with AS, and because they did not administer tests of theory of mind, their explanation must be viewed as speculative.

Pijnacker et al have opened up the tantalising possibility of distinguishing among the various levels of pragmatics and among people with ASD: implicature, on their view, need not be seen as challenging for people with ASD *per se*, but rather as challenging for people with HFA, who rely on compensatory strategies in verbal intelligence. While their findings prompt future research to focus on the implicature in autism, Pijnacker et al's account is limited in two respects that will prove important later in this thesis. One of their limitations is the operational definition of pragmatics, that is, their exclusive use of scalar implicatures. Pijnacker et al quite reasonably used scalar implicatures as a relatively neat subset of implicature and as a phenomenon about which there is much background experimental-pragmatic data. But Pijnacker et al also took scalar implicatures to be representative of implicature more generally, an assumption that may not hold. The point is not quite that scalar implicatures are not (or not wholly) implicatures, although that case has been made (e.g. Chierchia, Fox & Spector, *forthcoming*). The point is rather that even if scalar implicatures are fully inferential, they are more constrained than particularised conversational implicatures more generally, and may therefore become 'routinised' and hence, in processing terms, trivial. Pijnacker et al's study needs to be complemented with studies of less constrained implicatures if we are to draw a detailed picture of implicature in autism. The second limitation is in the way Pijnacker et al (*ibid*) relate their findings to the theory of mind. The relationship between theory of mind and pragmatics, as later chapters will argue, is contentious, with some researchers arguing that theory of mind is superfluous even for high-

level pragmatics such as implicature (e.g. Ruth Kempson, *personal communication*). An explanation of participants' performance in theory-of-mind terms, then, is potentially problematic.

Summary

This chapter has argued that there is mixed evidence in the literature for pragmatic dysfunction in autistic spectrum disorders. The best evidence for impairment is with deictic expressions, and there is evidence for subtle problems in disambiguation, hence there is evidence for impairment in level (1) pragmatics. However, it is unclear whether these problems are peculiar to people with autistic spectrum disorders or to people who also have language impairment. There are hints of impairment to understanding ad hoc concepts, hence to level (2) pragmatics; and also to use of sub-sentences, hence to level (3) pragmatics. There is no clear evidence of impairment to implicature, hence to level (4) pragmatics.

The next chapter will use the framework of pragmatics developed in this chapter against a corpus of conversational data.

Chapter 2: Linguistic data and theories of autism

Introduction

This chapter reports a naturalistic study on a corpus of clinical-pragmatic language. The chapter has multiple aims. Firstly, it gives evidence of impairment at the different levels of pragmatics. Secondly, it discusses how these impairments should be explained; it argues against explanations from the theories of autism, and introduces instead a number of linguistic alternatives. Thirdly, it offers a rationale for the later experimental chapters.

This chapter has the following structure:

1. Introduction to the corpus
2. Method: data extraction
3. Analysis and discussion of the examples
    i. What are the levels of impairment?
    ii. How can they be explained?
        a. Ambiguity and theories of autism
        b. Ad hoc concepts
        c. Phrasal idioms and Graded Salience
        d. Structural Impairment
    iii. Why experiment?

Introduction to a corpus of clinical language

While naturalistic data have a marginal place in experimental psychology, a naturalistic study can be beneficial for a research program such as the project reported in this thesis. Many studies on language and autism take large samples of children from clinical and control groups, and have the samples perform standardised tests to estimate the language abilities of the relevant clinical

populations. Such studies are invaluable but - as Kjelgaard and Tager-Flusberg (2001:305) note – they can only be used to grossly generalise about linguistic phenomena, because they do not allow fine-grained analysis of problematic linguistic structures. Naturalistic studies are an invaluable first step towards specifying an inventory of language problems because they allow hypotheses to be generated that can later be teased apart experimentally.

Although naturalistic studies' value is constrained, there are various strategies to mitigate their drawbacks. Any naturalistic study of autism is hostage to the fact that, despite language problems being characteristic, indeed diagnostic, of autistic spectrum disorders, these problems may not surface in the text. To counteract this limitation, naturalistic studies can employ two strategies, which are reflected in the corpus study. Firstly, a naturalistic study can elicit natural language in semi-structured conversation that probe problem areas. The corpus study reflects this strategy by including conversational data from Bishop and Adams (1992) and Bishop, Chan, Adams, Hartley and Weir (2000). The authors held semi-structured interviews with children with pragmatic language impairment (a disorder, like autism, on the spectrum of pervasive developmental disorders), children with specific language impairment, typically developing aged-matched controls, and typically developing younger controls. The interviewer used sequences of pictures as a springboard for unscripted conversations. The resulting data show important language problems in children with specific or pragmatic language impairment, as measured by standard psychological measures such as mean length of utterance. Because the data comprise accurate transcripts, they also allow the finer-grained linguistic analysis that will be outlined below. But this first strategy is, of course,

limited by the responsiveness of clinical participants: many of these participants simply produced too little language for sensible analysis. Accordingly, this corpus study includes a second strategy: namely, admitting anecdotal evidence. Much of this anecdotal evidence was collected from various researchers and clinicians. The rest of the evidence was extracted from a database run by Professor David Skuse at the Behavioural Sciences Unit of the Institute of Child Health, University College London; this database comprises parents' responses to questions on the Children's Communication Checklist (Bishop, 1998). With these two strategies, we can glean valuable evidence from naturalistic studies.

The naturalistic data on which this corpus study draws yield vivid, memorable examples of language dysfunction; however, the data are uncontrolled in any strict sense. The data are tagged for all errors and dys-fluencies for the clinical-pragmatic group alongside the typical and SLI control group from Bishop and Adams (1992) and Bishop et al (2000). The data were controlled in these original studies because the measures were such that controlled comparisons could be made: for example, the authors could investigate whether mean length of utterance differed between experimental and control groups. But for present purposes the data are less controlled: if an error is present in the clinical-pragmatic group but not in the control groups, this might be because the relevant linguistic construction happened not to surface in the control groups' conversation, and not because the control groups would not have made the mistake. In any case it is accepted practice (see, for example, Perkins et al, 2006) that, as long as researchers draw only cautious conclusions, they can use naturalistic studies as credible sources of evidence on autistic language.

There are certain assumptions that will be made to allow us to use the corpus as evidence. Firstly, we will assume that the reported errors occurred in similar contexts, or at least that important differences between contexts were averaged out. We can surely imagine that pragmatic dysfunction might be caused, not by problems with pragmatic competence *per se*, but by performance factors such as a particularly demanding speech context. To an extent, however, these problems are mitigated by the complementary data sets. One data set, the responses to the Children's Communication Checklist, offered only fragmentary contexts; but another set, the data from Bishop and Adams (1992) and Bishop et al (2000), offer exactly transcribed speech contexts; and the final set, donations from researchers and clinicians, typically came from detailed notes that were made at the time of the utterance in question. Secondly, we will assume that three distinct disorders - autism, Asperger Syndrome and pragmatic language impairment – are, in some relevant respect, homogeneous: at least for autism and Asperger Syndrome, this assumption is in keeping with recent APA opinion (APA, 2010). Against this assumption stand important diagnostic differences. For example, pragmatic language impairment does not entail restricted, repetitive autistic behaviours (Cummings, 2008). The disorders are unified, however, in that they all supposedly implicate pragmatics (*ibid*). Deficits in one of these disorders will be taken, then, to be relevant for the category of Autistic Disorder. But since the sources of the data drew on existing DSM guidelines, the formal diagnosis is given wherever possible.

With the assumptions outlined above, we can draw cautious conclusions and hypotheses for future research.

Method: Data Extraction

The examples reported here are from clinical-pragmatic speakers and are drawn from a streamlined corpus of some 125 problematic sentences extracted from the larger corpus described above; this streamlined corpus is included in the Appendix. To identify anomalies, expressions were first tagged for the following linguistic phenomena: indexical or context-sensitive expressions (personal, spatial, and temporal deixis; anaphora; ambiguity; phrasal idioms); contextually calibrated meanings (ad hoc concepts); pragmatic enrichment of minimal propositions; implicature (including metaphor and irony). Tagged examples were checked against the researcher's intuition, the researcher being a native speaker of British English; anomalies and dys-fluencies were compiled for clinical-pragmatic and control groups.

Analysis and Discussion of the corpus

(1) What are the levels of impairment?

The following examples illustrate the possible levels of pragmatic impairment in the clinical-pragmatic group.

2(a) Adult: Is your mum having a baby?
Child (high-functioning autism; henceforth 'HFA'): No, but she will in August.

2(b) Adult: Can you stand [bear] to do more?
Child (HFA): [stands up]

2(c) Adult: What's wrong with the boy?
Child (PLI): Hurting

2(d) Adult: Do your shoes slip up and down when you walk?
Child (HFA): No, they slip down and up

In example 2(a) the child is confounded by aspect. He misinterprets the adult's utterance. That utterance picks out an event, the mother's pregnancy, which covers the (approximately) 9 months from conception to birth. At any time during these nine months, the child should accept the statement, 'Your mother is having a baby'. But in replying the child picks out a different event: the punctual, or at least rather less protracted, event of giving birth: he has failed to assign a semantic value to a context-sensitive expression. This child is dysfunctional at Level 1.

In example 2(b) the child incorrectly disambiguates the word 'stand'. On hearing the word in context, the child should have understood by it, 'bear to do' or 'put up with doing', but instead understood by it, 'stand up'. Arguably then this child is likewise dysfunctional at Level 1. But alternatively the child in example 2(b) might be, not failing to disambiguate correctly, but rather failing to calibrate underspecified meanings to a context through ad hoc concept construction. Such examples will be discussed further in section 3.ii.b below; it suffices here to note a possible impairment to Level 2.

In example 2(c) the child makes a strange response, but any error is more marginal than in the previous examples. The child may not be completely wrong, but seems

to have omitted too much information for an acceptable sentence. Compare his response with a typical response by a control:

2(e) Adult: What are they building?
Control: A sandcastle

The control's answer is brief but acceptable; he omits information but replies with a constituent, and the elliptical elements 'they', 'are' and 'building' are easily recoverable from the immediate linguistic context. In contrast, although the clinical-pragmatic child in example 2(c) omits a similar amount of information and replies with a constituent, his elliptical elements, 'he' and 'is', are unrecoverable from the linguistic context. The clinical-pragmatic child's ellipsis is prompted by 'what is wrong with the boy', for which a noun such as 'a headache' could stand, but for which the gerund 'hurting' could not. If the child had responded, for example, with 'a headache' (or arguably even just 'headache') then his response would have sounded natural. The following two examples from other children in the corpus suggest a superficially similar deficit:

2(f) Child (PLI): I [am] bigger than the other two.
2(g) Child (PLI): There's a policeman [who] doesn't look after him a lot.

But these examples are more obviously structural than example 2(c), because the elliptical elements are not full constituents; example 2(c) might have a less structural cause in that the clinical-pragmatic child may have over-generalized an acceptable operation. The child could have meant his response as what Stainton (2004) has called a 'non-sentence'. Non-sentences are utterances that resemble ellipsis, but that, unlike ellipsis, are mandated by the non-linguistic context. An example non-sentence is, 'fast relief from arthritic pain' written on the side of a

medicine bottle: there is no linguistic context; the sentence fragment is acceptable because of the non-linguistic context. Another example non-sentence is, ‘Sam’s mum’ said by a speaker as a woman enters the room: the non-sentence successfully conveys the meaning, ‘That is Sam’s mum’ but is again mandated by the non-linguistic context. Non-sentences could in principle be semantic ellipsis, where a single word conventionally has an elliptical content (e.g. ‘Cheers!’ means ‘Enjoy your drink!’), but then cases such as ‘Sam’s mum’ would have to have a conventional alternative meaning of ‘That is Sam’s mum’, which seems implausible. These non-sentences must, on Stainton’s view, be decoded and fleshed out, using non-linguistic context, to create a full proposition. Because non-sentences fit best with Level 3 Pragmatics, the clinical-pragmatic child in example 2(c) shows equivocal evidence for impairment to Level 3.

In example 2(d) the child makes a striking response; his response is certainly unusual and inappropriate, but is it pragmatically dysfunctional? The child has clearly misunderstood the adult’s question but in so doing has recovered a conventional implicatum: ‘and’ is said to conventionally implicate ‘and then’; he rejects the adult’s utterance because of the implied sequence, which is false. The child has not understood that ‘up and down’ function as a symbiotic pair - a fixed but non-figurative idiom (e.g. also ‘fish and chips’ rather than ‘chips and fish’). A symbiotic pair does not carry the normal temporal implications: if I say, ‘I walked up and down lots of stairs today’ I don’t necessarily imply that I always started at the bottom of a staircase, walked up it, and then down again. Example 2(d) is the only problematic sentence involving implicature, and strikingly the implicature was

not the source of the problem. There is no evidence in the corpus for impairment to Level 4 pragmatics[12].

To summarize, the corpus offers possible evidence of pragmatic dysfunction at three different levels. This raises the question of how this apparent pragmatic dysfunction might be explained. Various explanations are considered in the following sections, which cite further examples from the corpus.

(2) Ambiguity and theories of autism: a test case

The corpus shows problems with disambiguation. The following examples are all taken from children with high-functioning autism.

2(e) Adult: Can you **stand** to do some more?
Child (HFA): [stands up]

2(f) Adult: How does that **look**?
Child (HFA): By its eyes.

2(g) Adult: (to child who is cutting with scissors) **Watch** your hands!
Child (HFA): [stares fixedly at hands]

2(h) Adult: Go and get yourself a drink from the **toilet**.
Child (HFA): [drinks from the toilet]

2(i) Adult: Mind if I **join** you?
Child (HFA): With this? [offers the speaker adhesive tape]

[12] Although we could argue that enriching 'and' to 'and then' is a more basic process than implicature, we could still conclude that there is no evidence in the corpus for impairment to implicature; accordingly we need not pursue such a theoretical question in this thesis.

The misinterpretations in examples 2(e) to (h) implicate words with multiple meanings. Each misinterpretation seems to result from a child choosing a contextually inappropriate meaning of the bold-faced word.

These examples can be described in more linguistic detail. In each example, the child with high-functioning autism chooses a contextually inappropriate meaning of the bold-faced word. Also, in each of these examples, the child chose a more concrete sense than the one intended. In examples 2(e) to (g) the children also recovered a different thematic structure from the one intended. In 2(e) the child should have taken his role to be the experiencer of a certain mental event, namely the enduring of an activity, but he instead took his role to be the agent of an activity, namely standing up. In 2(f) the child should have taken the animal to be the theme of an event, but he instead took the animal to be the agent of an action. And in 2(g) the child should have taken his role to be the experiencer of the mental event of attending to his hands, but he instead took his role to be the agent of a physical event. Moreover, in each example above the children cleaved to the dominant meaning[13] of the ambiguous word, that is, to its most common meaning. Only one example in the corpus did not fit this trend. In this example, the child recovered a subordinate meaning of the word 'island':

[13] Two points are due here. Firstly, meaning dominance was established in a free-association task; 6 raters unanimously identified dominant meanings for the ambiguous words that were the same as the meanings recovered by the children. Secondly, meaning dominance's relationship with such features as concreteness is unclear: it is unclear whether dominance itself is directly reflected in the mental lexicon or whether it arises out of various lower-level features. That is, is dominance a primitive, non-analyzable property, or is it a function of features such as concreteness or familiarity; is it just meanings themselves that are dominant or thematic structures?

2(j) Child (HFA): Robinson Crusoe was marooned on an **island** because he got half way across the road and he was too nervous to go the rest of the way

How might these examples fit with mainstream theories of autism? This section will now use disambiguation to test how well theories of autism can explain apparent pragmatic dysfunction. The three major theories of autism are outlined and applied to the data in turn.

The Theory-of-Mind Hypothesis originates in work by Baron-Cohen, Leslie and Frith (1985), who identified a deficit shown by children with autism in attributing mental states to others and in predicting and interpreting people's behaviour in light of these mental states. This mentalistic deficit is often discussed in terms of meta-representation. Meta-representations, according to Leslie and Roth (1993) for instance, are representations that comprise an agent, a propositional attitude, a primary representation and an imagined or decoupled representation. For example:

2(k) A mother (agent) pretends (propositional attitude) of the banana (primary representation) that it is a telephone (decoupled representation).

(Leslie & Roth, 1993)

Problems forming such representations are supposed to explain autistic deficits in standardized tests of theory of mind, such as false-belief tasks and the strange-stories task. A meta-representation such as 2(k) is first-order; first-order representations can be embedded within second-order meta-representations, as, for example, in the following:

> 2(l) [2nd-order meta-representation A son (agent) believes (propositional attitude) that [1st-order meta-representation his mother (agent) is pretending (propositional attitude) of the banana (primary representation) that it is a telephone (decoupled representation)]].

Impairment to meta-representation can occur just with second-order meta-representation or with both first- and second-order meta-representation.

Predictions from theory-of-mind differ widely with respect to language. Some predictions take the theory loosely: Martin and McDonald (2003), for instance, predict deficits in language relating to social and emotional domains. On this loose view, there is a set of isolable expressions that refer to social and emotional phenomena. This view is consistent with deficits in mental-state vocabulary (Tager-Flusberg, 1993) and with examples 2(e) to (g) and 2(i). Other predictions take the theory more strictly: deficits relate to an underlying capacity for meta-representation. On this view, deficits will show up whenever meta-representation supports language, whether it is explicitly social or emotional, or implicitly mentalistic. This view is consistent with distinctions among similes, metaphors and irony in speakers with autism: similes are supposed to require no meta-representation; metaphors, first-order meta-representation; and irony, second-order meta-representation (Happé, 1993; though see Norbury, 2005, for some qualifications).

The strict Theory-of-Mind Hypothesis, based on the meta-representation framework, can be put to work in explaining problems with disambiguation. Such an account will focus on contextual inferences rather than underlying lexical features. Mentalistic inferences, we could argue, are needed for disambiguation, because they allow a listener to work out what the speaker intended to say, and

thereby to resolve the ambiguity. If such first-order meta-representations are impossible, or at least very difficult, for some children with autism, then they should not be able to resolve ambiguities. But children and adolescents with autism can access multiple meanings of ambiguous words with and without context (Norbury, 2005)[14]. A more moderate mentalistic deficit, we might counter, could instead explain problems with disambiguation; but such a deficit would surely still predict very wide-ranging difficulties, and Norbury (2005) found only small-scale difficulties. For the Theory-of-Mind Hypothesis to be put to effective use, then, we would need evidence that underlying these apparent autistic abilities is a selective disruption to those cases of disambiguation which most call on knowledge of a speaker's mental states. There does not appear to be clear evidence in the literature for such impairment. A theory-of-mind explanation therefore has limited appeal in the case of problems with disambiguation.

There is an alternative, more fundamental challenge to the use of theory of mind in explaining pragmatic dysfunction; this challenge will be taken up in the following chapter.

Weak Central Coherence (WCC) seeks to explain deficits that, though apparent in autistic spectrum disorders, cannot readily be explained by the Theory of Mind Hypothesis. Examples of such deficits are the following long-established findings: that participants with autism remember word strings well but not sentences

[14] Norbury (2005) did not administer theory-of-mind tasks to her children, so her data do not show whether her children were impaired at first- or second-order theory of mind; but her data nevertheless suggest that strict theory-of-mind predictions, as defined above, are too strong.

(Hermelin & O'Connor, 1967) or that readers with autistic spectrum disorders tend not to select semantically appropriate pronunciations of homographs (Frith & Snowling, 1983). Weak Central Coherence purports to explain such surface deficits not so much by invoking an underlying cognitive deficit as by invoking a different style of cognition and perception. In this weakly centrally-coherent processing, processors make limited use of context and, as a result, prefer local processing to global processing: local detail, such as individual words, is favoured rather than global detail, such as sentence context. Weak Central Coherence has been applied to autism but also to William's Syndrome, another developmental disorder which brings about unusual linguistic and cognitive skills.

WCC is immediately appealing in the case of disambiguation in autism. For there is similarity between disambiguation and some early motivating evidence for WCC, namely reading the correct pronunciation of a homograph. WCC would explain a deficit in disambiguation, such as that seen in the examples above, roughly as follows. WCC causes participants with ASD to focus on local processing and neglect the global context; WCC, that is, causes participants to neglect the context and to default to the context-free meaning of the words, i.e. the dominant meaning.

Although the WCC account is attractive, it may not hold. For when the WCC notions of local and global processing are considered further, problems with them rapidly emerge and the question arises of how local processing and global processing speak to the notion of context. In a sense, context does not seem badly affected in these examples: in each case, a child responds in a way that is grammatically correct. If a child heard, 'Can you stand to do some more?' he could

respond perfectly acceptably by standing up given the right context, say, if the child was being asked to do some more physical exercises after a break. It is not the grammatical correctness that is at question here but the appropriateness to the speech context.

So the question arises of how WCC relates to non-linguistic context. Because a non-linguistic context comprises information of many forms and from different modalities, non-linguistic context presumably demands higher-level processing than does sentence context. But non-linguistic contexts and their high level of processing might be incompatible with WCC; for in WCC high- and low-level processing are in competition (Martin & McDonald, 2003). High-level processing is favoured in WCC-autism research, where data is amassed for such tasks as sentence completion. But recall that Central Coherence is also applied to William's Syndrome. Low-level processing is favoured in WCC-Williams-Syndrome research, where data is amassed for basic perceptual tasks. A holistic WCC would need an autistic local-processing bias in low-level processing, which is hotly disputed (Happé, 1996; Rinehart et al, 2000; Ropar & Mitchell, 1999), and a Williams-Syndrome local-bias in high-level processing, which is similarly controversial. In favour of such a bias in high-level processing, some Williams-Syndrome researchers have suggested deficits such as a delay in lexical acquisition underpinned by a bias away from semantics and towards phonology (Grant, Karmiloff-Smith, Gathercole, Paterson, Howlin, Davies & Udwin, 1997), a possible deficit in the syntax of relative clauses (Grant, Valian and Karmiloff-Smith (2002) and abnormal pragmatics triggered by stereotyped responses (Laws & Bishop, 2004), Against such a bias, other Williams-Syndrome researchers have suggested that language and mentalistic skills are within

the normal range (Karmiloff-Smith, Klima, Bellugi, Grant & Baron-Cohen, 1995). Whether WCC predicts impairment at such high levels as compositional semantics or pragmatics in processing non-linguistic contexts is unclear. Accordingly is unclear whether deficits in disambiguation of the kind shown in these examples are evidence for WCC.

Executive Dysfunction seeks to explain different data from both the Theory of Mind Hypothesis and WCC. Executive-dysfunction researchers have identified an autistic deficit in suppressing prepotent responses in tasks that require a combination of working memory and inhibition (Diamond, Prevor, Callender & Druin, 1997; Roberts & Pennington, 1996). Executive Dysfunction draws evidence from such cases as a participant with autistic spectrum disorder persevering inappropriately with a previously correct rule in the Wisconsin Card Sorting Task, or with a direct-route rather than a detour-reaching route in Hughes and Russell's' (1993) box task, or with a previously correct but now inappropriate pronunciation in a homograph task (Hala, Pexman & Glenwright, 2007).

Executive Dysfunction, in contrast with the Theory of Mind Hypothesis, fits better with lexical dysfunction than inferential dysfunction, and executive dysfunction can plausibly explain problems with disambiguation. Executive functions are likely to be needed in recovering the intended meaning of an ambiguous word. To exemplify the executive demands, consider the following: a hearer must access the lexical entries for a given word, hold them in mind, and act on them to select the contextually appropriate sense; but in many cases, such as when a dominant (or prepotent) meaning is inappropriate, a hearer will also have to suppress an

unintended prepotent response. Inability to suppress the prepotent response might well explain the examples above. In example (9), for instance, a hearer would have to inhibit the dominant, hence prepotent, meaning of 'look', the action of looking, to access the subordinate meaning 'seem'.

Executive Dysfunction gives a plausible account of cases where there is a pre-existing set of candidate meanings, but where such sets do not exist the theory meets difficulties. There may, for example, be no such sets when a word is radically contextually calibrated: witness the many interpretations of the word 'fresh' discussed in the introduction (Murphy, 1997). A possible example in the corpus is this:

2(m) Adult: **Put** your **milk** on the table.
Child (HFA): [pours milk on table]

Here the adult arguably loosens the concept MILK so that it includes the container holding the milk[15]; she does not simply select the meaning from the lexicon. Such cases raise problems for Executive Dysfunction in the following way. Executive Dysfunction is typically assessed with respect to a set of pre-existing possibilities: for instance, a prepotent response in the context of Hughes and Russell's Box Task is a direct route rather than a detour-reaching route (Hughes & Russell, 1993). But with ad hoc concepts the options do not already exist; they must rather be created in context. Nor is it quite the case that the strict, literal concept must be straightforwardly inhibited; the concept must, instead, be worked on and adapted to the context. It is not obvious how to interpret these points with respect to Executive

[15] More traditionally we might call this metonymy, but that term is descriptive and not explanatory; ad hoc concept construction, as will be discussed later, provides an explanation.

Dysfunction. Does a strict concept qualify as a prepotent response? The answer is unclear. If it does, it is a rather different form from the usual definition. Ad hoc concept construction will be revisited in 3.ii.b as an alternative to the theories of autism.

There is, in any case, a question about how – and whether -Executive Dysfunction should apply to language at all. As Joseph et al (2005) remark, 'the different possible links between language and executive impairment in autism have not yet been clearly conceptualized and evaluated in relation to one another'. Given this lack of clarity it is hard to set Executive Dysfunction to work in explaining language problems. Indeed the theory may not, as Joseph et al (*ibid*) suggest, make predictions for language. These authors searched for an association between tests of language and of executive function in children with autism or pervasive developmental disorder-not otherwise specified (PD-NOS) and in typically developing children. While they found evidence of an association between the two sets of tests in typically developing children, they found no evidence of such an association in children with autism or PD-NOS. They concluded, accordingly, that executive dysfunction and language impairment in autism are not directly linked. Instead, they suggested that children use language to facilitate their performance in executive function tasks – say, for example, through inner speech - but that children with autism or PD-NOS are unable to do so. If Executive Dysfunction does indeed only apply to this special function of language, then the theory may not have any direct predictions at all about the traditional levels of linguistic representation.

As we have seen with the two preceding theories, though, there are multiple ways to interpret Executive Dysfunction. On the narrower sense - an inability to suppress prepotent responses in tasks involving working memory and inhibition – there may not be clear links with language, especially in phenomena such as *ad hoc* concepts. But there is a broader sense in which Executive Dysfunction comprises any impairment to executive functions. On this sense, there is clearer evidence of an association with language. For instance, Bishop and Norbury (2005) showed an association between difficulties in tasks assessing generativity and communicative impairment. Generativity tasks require the participant to generate novel ideas. Most sensitive to autism are tests of ideational fluency. In these tests a participant might, for example, suggest novel uses for an object or interpretations of a meaningless pattern. These more creative tasks bear some similarity to the kind of creativity required in generating and interpreting *ad hoc* concepts. It is unclear, however, whether we should broaden Executive Dysfunction to include other components of executive function, and thereby lose the specificity that obtains for the theory.

By pressing the three major theories of autism into service in explaining problems with disambiguation, we can see their limitations. Theory of Mind risks predicting too much if it focuses on meta-representation and too little if it focuses on social, mental and affective language. Weak Central Coherence is subject to uncertainty about what levels of linguistic representation it should apply to. Executive Dysfunction seems limited to cases when there are pre-existing candidate sets and, in any case, may not be appropriate for application directly to language, as argued by Joseph et al (2005). In all cases, and especially WCC and Executive Dysfunction, there are numerous possible interpretations of the theories of autism.

To some extent, problems with the theories are mitigated by this fact, because researchers can choose their own interpretations of the big theories and test these interpretations. If we regard the theories as being useful perspectives on autism, rather than concrete theories, then we can interpret them in this way with impunity. This is an acceptable, top-down strategy. An alternative approach is to take a bottom-up strategy: that is, we can focus instead on particular deficits and seek local (in this case, linguistic) explanations and afterwards relate them to the theories of autism. The next sections suggest how such linguistic explanations might run.

(3) Ad Hoc Concept Construction

In discussing ad hoc concepts, the preceding section hinted at a first linguistic alternative to the theories of autism. An ad hoc concept, as we have seen, is an instance of a lexical item that has been calibrated to a certain context. Ad hoc concepts can be broader than the encoded concept, as in 'milk' for 'glass with milk in'; they can be narrower, as in 'drink' for 'alcoholic drink' (e.g., 'I really need a drink'); and they can be metaphorically extended, as in 'bulldozer' for a forceful person (Wilson & Carston, 2007). Ad hoc concepts have been given a persuasive linguistic explanation by researchers working in Relevance Theory (e.g. Wilson & Carston, *ibid*). This explanation will be outlined here, in brief, to show that linguistic accounts can offer potentially valuable explanations of autistic deficits.

The following account is due, principally, to Wilson and Carston (*ibid*). The theory of ad hoc concept construction draws on established examples of broadening, narrowing, and metaphorical extension, and on Barsalou's (1993) work in problem solving and reasoning. The theory centres on the suggestion of a common process that operates on a range of possible meanings, from literal meanings through to loose uses such as broadening and narrowing, and finally to metaphor and hyperbole. Underlying these phenomena, relevance theorists suggest, is a cost-benefit style processor. Relevance theorists give the following account of this processor. Relevance theorists suggest that cognition is geared towards a search for relevance. An input to cognition is relevant if processing it in context yields a positive cognitive effect: for example, a true conclusion that could be drawn only from input and context together (Wilson & Sperber, 2002). Relevance comes in degrees: humans attend to inputs that are the most relevant, that is, from which they can derive the most cognitive effects. To borrow an example from Wilson and Sperber (2002), one's train arriving half an hour late is, under normal circumstances, more relevant than one's train arriving one minute late, because the latest train will prompt more cognitive effects regarding the necessary reorganization of one's schedule. Relevance is not solely determined by cognitive effects, however; it is also a function of processing effort. In other words, the more effortful it is to process an input, the less relevant that input is. To borrow an example from Van der Henst and Sperber (2004), if one wants to catch the next train to Bordeaux, then the input, 'The next train to Bordeaux is at 3:24pm' will be more relevant than, 'The next train to Bordeaux is at 36 minutes before 4pm', because while the information content is the same, that content is easier to recover from the first input than the second. According to Relevance Theory, humans

follow a path of least effort and fix on the first interpretation that satisfies their expectation of relevance (Wilson & Sperber, op cit). Relevance-theoretic pragmatics is the communicative reflex of these general cognitive principles: our contextual inferences from coded meanings to intended, implicit meanings operate along the very same lines.

With regard to *ad hoc* concept construction, the phenomena of narrowing and broadening are triggered by a search for relevance and follow a path of least effort; the phenomena stop when expectations of relevance are satisfied. Examples of metaphorical extension, such as calling a man a bulldozer or a woman a princess, are said to result from processes of narrowing and broadening operating in tandem. Thus in metaphorical extensions there are emergent features: a man who is a bulldozer is insensitive or aggressive, but no actual bulldozer can be insensitive, in the relevant sense, or aggressive. Although the relevance-theoretic view of *ad hoc* concepts has not been subjected to experimental tests, there is no reason why this view should not be so tested. There is potential here to unify under ad hoc concept construction various linguistic phenomena and non-linguistic phenomena. For example, perhaps ad hoc concept construction, with Relevance Theory, could explain performance on tests of ideational fluency.

Although this section has been speculative, it serves to illustrate how useful a bottom-up (i.e. linguistics-driven) approach can be. The following section will draw on another problematic area in the corpus, namely phrasal idiom comprehension, to sketch a further possible linguistic explanation.

(4) Phrasal Idioms and Graded Salience

Phrasal idioms are traditionally defined as phrases whose meanings are somewhat fixed and figurative and do not depend on a strict composition of the meaning of their parts. Phrasal idioms are stereotypically conventional, in the sense that there is a mismatch between any literal meaning that the phrase may have and the figurative meaning; opaque, in the sense that the motivation for the figuration underlying the idiom is unclear; and non-compositional, in the sense that the figurative meaning cannot be decomposed onto the parts of the phrase (Nunberg, Sag & Wasow, 1994). An idiom that exemplifies all these features is 'to kick the bucket'. Phrasal idioms are not, though, a natural kind; they vary in the extent to which they exemplify the features listed. In their seminal paper on the subject, Nunberg, Sag and Wasow (1994) posited two categories that differ around what they view as a pivotal feature, namely, their compositionality. The first category is phrasal idioms: these may vary in some features but are non-compositional, in the sense that the meaning cannot be decomposed on to the idiom's parts. Thus, 'kick the bucket' (i.e. to die) and 'saw logs' (i.e. to snore), which are figurative, are phrasal idioms, but so too is 'by dint of', which is not figurative. The second category is idiomatic combining expressions: these expressions are idiomatic in virtue of their conventionality and figuration, but the meaning of the idiom can be decomposed onto the meaning of its parts. Thus, 'take advantage', 'pull strings' (i.e. to draw on one's influence to some specific goal) and 'spill the beans' (i.e. to divulge a secret) are idiomatic combining expressions. These expressions are somewhat freer in structure: they can be modified, as in 'leave no legal stone unturned'; parts of them can be topicalized, as

in ‘those strings, he wouldn’t pull for you’; and parts of them can be ellipsed, as in ‘my goose is cooked but yours isn’t’.

The corpus reveals various examples of problems interpreting idiomatic expressions. Consider example 2(m) (example 2(d) above).

2(m) Adult: Do your shoes slip up and down when you walk?
Child (HFA): No, they slip down and up.

In this example the phrase ‘up and down’ is, as we saw above, a symbiotic pair. Symbiotic pairs are idiomatic in virtue of their fixedness although they are not conventional, figurative or opaque. They are, according to Nunberg, Sag and Wasow’s (1994) system, idiomatic combining expressions, because their meanings can be decomposed onto the parts of the expression yet the expressions have become fixed. To say that the shoes slipped ‘down and up’ sounds as odd as saying ‘The man ran down and up the street’; the order of ‘up and down’, in this case, does not imply that the ‘up’ action happened before the ‘down’ action.]

There are also examples, in the corpus, of problems with phrasal idioms.

2(n) Adult: It’s raining cats and dogs.
Child (HFA): [Reacts with deep distress]

This expression is a phrasal idiom, in Nunberg et al’s (1994) terms, because it is thoroughly non-compositional; it is also conventional, figurative and opaque. Indeed, in all of the characteristics that make the stereotypical idiomatic expression, example 2(n) is the converse of example 2(m). The child interprets the expression as though it were the literal, the compositional meaning. This is in spite of the fact

that a literal interpretation is nonsensical: there is no sense in which cats and dogs can literally rain from the sky. This fact ought to be a strong cue that an interpretation other than the compositional meaning of the expression was intended. An understanding of the real world seems either to be lacking or to be somehow isolated from utterance interpretation.

The following examples show further variety in the idiomatic expressions that are problematic for people with autism.

2(p) Adult: When are you taking off for Brisbane?
Adult with HFA: We're not taking off; we're driving.

2(q) Adult: We've tried recycling but it's got flies on it.
Adult with HFA: You could catch a disease.

Example 2(p) is compositional, conventional, figurative and non-opaque. Example 2(q) - which contains the Australian idiom 'it's got flies on it', an expression of disapproval - is non-compositional, conventional, figurative, and reasonably non-opaque. In both these examples, the problem is, again, that the hearer with ASD has recovered a compositional meaning of the phrase. Witness, also, examples 2(r) and 2(s):

2(r) Adult: You've got a little while left, Gina.
Child (HFA): Gina want (sic) a thousand whiles left. Gina likes the whiles at here.

2(s) Adult: Hop on the scales.
Child (HFA): [Hops onto the scales]

Example 2(r) is compositional, non-conventional, non-figurative and transparent; example 2(s) is compositional, conventional, figurative and transparent. These idioms, too, have been interpreted compositionally.

It is important to recognize that this tendency to be overly compositional is not unique to speaker-hearers with ASD. Some studies suggest that, until children are about 7 years old, they tend to prefer compositional meanings (Levorato & Cacciari, 1995; Ackerman, 1982). Remarkably though, some overly compositional interpretations come from an adult with high-functioning autism. This hints at, at least, a slow developmental trajectory, and perhaps even a more fundamental impairment. To explain such impairment, we could argue for a difficulty, in speaker-hearers with ASD, in forming complex constituents. There is precedent for similar deficits in language disorders: Van der Lely and Christian (2000) found that children with a subtype of SLI displayed abnormal behaviour when pluralizing compound nouns, offering evidence of difficulty in successfully building complex lexical units, albeit at a much more basic level than phrasal idioms. The possibility that speakers with ASD have a deficit in forming complex lexical units does not seem to have been directly tested, although it is suggestive that Norbury (2004) found sentence processing abilities to be a good predictor of idiom comprehension. An important qualification is due, however: speaker-hearers with ASD are not universally overly compositional in their interpretations of idiomatic expressions. Consider the following:

2(t) Child (HFA): I'm going to bed uncharacteristically early.

What is unusual about this example is that, since hearing the line in a film, the child uses it every night on going to bed, at whatever time. The child behaves as though she has created her own phrasal idiom. She has, in fact, inappropriately created a complex lexical item. This example is by no means isolated; it relates quite closely to the kind of stereotyped language reported in autistic spectrum disorders in which

phrases, and indeed intonation patterns on these phrases, become fixed. A simple deficit in forming complex lexical items, then, does not look promising, although controlled experimental testing would be worthwhile in assessing the possibility.

To master idiomatic expressions, speaker-hearers must be able to produce and recover complex lexical units; they must also be able to disambiguate them from the literal meanings of the expression. An intriguing possibility, then, is that problems understanding phrasal idioms are reducible to problems with disambiguation, albeit a more complex type of disambiguation than we saw above. We could proceed, here, by asking whether the three major theories of autism can explain deficits in idiom comprehension. In so doing, we would get no further than was already achieved in the section above, without, that is, advancing more formal models to implement the theories. Instead we will turn to a more formal model existing already in linguistics, a model that incorporates many of the features discussed in the section on the linguistic features of the disambiguation problems.

The graded-salience model (Giora, 1997, 2002; Peleg, Giora & Fein, 2004) posits parallel lexical and contextual processing. On this model, lexical processing and contextual processing are supported by distinct mechanisms. Lexical meanings are encoded and ordered with respect to salience; salience is a function of such lexical features as conventionality, concreteness, typicality, frequency and meaning dominance. Only encoded meanings can be salient. Encoded meanings are activated by the linguistic input, with salient meanings being prioritized over non-salient ones. Contextual processing affects utterance interpretation from the start in

the sense that it can predict meanings, but it cannot intervene in lexical processing and cannot block salient meanings. Thus, context will affect utterance interpretation only later on, and contextually inappropriate salient meanings will still be activated even with strong opposing contexts. Strong contexts can, however, speed up processing such that literal and figurative meanings can be accessed equally quickly. The leading advocate of graded salience has been Giora (e.g. Giora, 1997, 2002; Giora & Fein, 1999a; Peleg, Giora & Fein, 2004).

As far as phrasal idioms are concerned, graded salience makes a distinction between familiar and unfamiliar idioms: familiar idioms are salient; less familiar idioms will be less salient; unfamiliar idioms will have to be inferred. Familiar and literal meanings are, on this model, active in appropriate and inappropriate contexts; less familiar, in appropriate contexts only. Although experimental tests have focused on familiarity, Graded Salience also takes account of conventionality, concreteness and so on; thus Graded Salience can potentially distinguish among phrasal idioms on various different dimensions. Graded salience can thus make a range of predictions for phrasal idioms; essentially the same account is given as of ambiguous words. Deficits in disambiguation and in understanding phrasal idioms can be localized experimentally to lexical processing on the one hand and contextual processing on the other by manipulating salience and inference: Graded Salience makes specific reference to linguistic features such as those that the three major theories of autism do not speak to.

Graded Salience makes interesting predictions about the features implicated in the corpus, but these predictions are of limited use with naturalistic data. If an idiom is

familiar it will be more salient, hence easier to process; if it is concrete (there is no figuration) it will be more salient; if it is dominant relative to the literal string it will be more salient. But a property such as familiarity is only meaningful relative to the individual speaker or hearer. Granted that, in the corpus, the hearers did not understand the phrasal idioms in context; we cannot conclude that they were unfamiliar with the idioms. To manipulate familiarity, researchers must obtain estimates of familiarity from their participants, for example by pre-testing. A property such as the dominance of the phrasal idiom relative to the literal string would likewise have to be estimated. In contrast, how concrete an idiom is can simply be read off the phrase itself; but in this case even a concrete idiom such as 'up and down' seems problematic for clinical-pragmatic participants. Moreover, to access an inferential component in phrasal-idiom comprehension, we would have to introduce unfamiliar idioms to participants in controlled contexts. Overall, then, Graded Salience can best be assessed experimentally.

It is important to stress, at this point, that linguistic approaches to language deficits need not be held in opposition to theories of autism. Linguistic explanations are proximal: they seek to explain deficits in terms of specifically linguistic competence and processing. Psychological explanations are distal: they seek to site language deficits in cognition more generally, and aim for a unified cognitive account. When viewed in this way, linguistic and psychological explanations are quite compatible; indeed, an account that incorporated both types of explanation would be powerful.

(5) Structural Impairment

So far we have tacitly assumed that if there are superficial impairments to the levels of pragmatics outlined in the previous chapter, then this is evidence for pragmatic dysfunction. This need not be the case. The graded-salience viewpoint sketched above, for instance, does have a pragmatic-inferential component, but it also has a strong lexical component. This raises the question of whether deficits in clinical-pragmatic speaker-hearers might better be characterized as structural rather than pragmatic.

We return to the corpus. Consider the following examples:

2(u) Child (PLI): Them are a bit dangerous

2(v) Adult: Have you ever ridden a horse?
Child (HFA): Yes, it's always itchy now.

2(w) Child (PLI): It could be Fred. (Fred was not present in discourse)

These are all examples of person deixis. The examples above illustrate both pragmatic and structural aspects of deixis. Witness example 2(u): in this example, 'them' is used instead of the grammatically correct 'they'. This seems therefore to be a structural problem rather than a pragmatic one. Case has been incorrectly assigned: non-nominative case is used instead of nominative. This error could be due to incorrect assignment of case *per se* - hence to a morpho-syntactic deficit - or to difficulties in acquisition leading to an incorrect lexical entry for 'them', hence a

semantic deficit. Example 2(v) is less clearly structural: it is an instance of what Fine et al (1994; see review chapter for discussion) have called an unclear or additioning reference, which is to say there is no apparent antecedent fot 'it' to refer back to. The mind is left to boggle at what is itching. Likewise, with example 2(w), the proper noun 'Fred' is used without a preceding reference to Fred; in this context, the child in fact named a character *ad hoc* in a picture that he was describing. We are left with a mixed picture with regard to deixis: clinical-pragmatic children show both deficits that are plausibly structural and deficits that are plausibly pragmatic.

There is a similar picture for examples of abnormal article use. While article use – and indeed definiteness and indefiniteness more generally – resists neat classification (Abbott, 2004), there are intuitions about usage that allow us to compare clinical and typical uses. Definite articles tend to imply uniqueness, familiarity or specificity; indefinite articles, the converse (*ibid*). Bearing this rough distinction in mind, consider the following examples.

2(x) Child (PLI): He's in different class.

2(y) Adult: How did you get there?
Child (PLI): In the aeroplane.

2(z) Child (PLI): Took a picture where a man went flying. A man died.

In example 2(x) the child incorrectly omits the article: an indefinite article is required. In example 2(y) the child uses the definite article when the indefinite would have been better. This is a subtle abnormality which is made clearer by the fact that no aeroplane was mentioned in the preceding discourse; typically, when introducing an unfamiliar entity, in the sense of the entity having not appeared in the preceding discourse, the indefinite article would be used (*ibid*). Example 2(z)

shows the reverse of this problem: an indefinite article is used instead of a definite one. Because the second occurrence of 'A man' refers to the same person (this is made clear by the broader speech context), the definite article would be better. Using the definite article to refer to familiar entities in this way improves the coherence of the discourse by encouraging the hearer to interpret the discourse as a single unit rather than disconnected sentences (Gernsbacher & Robertson, 2001). Because the child in example 2(z) uses the indefinite article, we are encouraged to view the sentences as disconnected and cannot tell whether the child himself intended the sentences to be related. It is unclear whether such subtle problems are semantic or pragmatic; there is indeed much debate on the subject (Abbott, 2004). Example 2(y) looks quite likely to be a structural problem in that the article is altogether absent. Good cases can be made for examples 2(y) and 2(z) to be either semantic, if a more dynamic semantics is adopted (e.g. Heim, 1983), or pragmatic (Abbott, 2004). Thus we must remain agnostic as to a pragmatic impairment here.

This section will draw to a close by remarking that there are a few cases, in the corpus, of seemingly unambiguous structural impairment. Note the following problems with preposition usage:

2(aa) Child (PLI): They're going in sand.

2(bb) Child (PLI): In the beach.

2(cc) Child (PLI): I've been to the holiday.

2(dd) Child (PLI): At 2nd February.

While we could argue that examples 2(aa) and 2(bb) show abnormal conceptualization of space, this would imply deficits in conceptual representation, hence in semantics.

This section has served to show an ambiguity in the data between pragmatic and structural impairments; only controlled experimentation can serve to tease apart these possible impairments.

Summary: Why experiment?

This naturalistic study has supported the argument developed in the literature review, that there are hints of impairment at levels (1), (2) and (3) but a lack of evidence for impairment at level (4). The study has also pointed out various weaknesses in the application of theories of autism to linguistic data, and various possible strengths in the application of more proximal linguistic alternatives. In order to support these naturalistic claims, a more controlled study is required. The experimental studies reported in the following chapter are an attempt to search, in a controlled context, for evidence for impairment to level (4) pragmatics, while using linguistic models as alternatives to the theories of autism.

Chapter 3: Introduction to Experiments

This chapter introduces experimental work carried out on particularized conversational implicature in typical adults, adults with Autistic Disorder, and typically developing children. Firstly, it lays out the phenomenon to be investigated - the implicature - and the justification for the investigation. Secondly, it discusses an experimental method that approaches the implicature orthogonally. Thirdly, it discusses problems with explaining pragmatics using theory of mind. Finally, it introduces a design, based on the experimental method discussed, that tests implicature more directly, avoiding various experimental and theoretical problems.

The Particularized Conversational Implicature

The General Introduction laid out a definition of pragmatic dysfunction and, implicitly, the following view of pragmatics:

(1) Pragmatics assigns semantic values to indexical and otherwise context-dependent expressions: for example, deictic expressions such as pronouns or tensed expressions.
(2) Pragmatics calibrates lexical items to the context: for example, *ad hoc* concepts such as context-sensitive uses of adjectives (e.g. 'fresh').
(3) Pragmatics saturates and freely enriches minimal propositions (e.g. 'I've had a very big breakfast [today]').
(4) Pragmatics derives implicatures (for present purposes, this includes irony and metaphor)

The experimental work in this thesis will focus on Level (4), the implicature, on the following grounds. The General Introduction outlined how various processes might support pragmatics. A good case has been made, for instance, for primary pragmatic

processes, which are sub-personal, sub-doxastic, non-inferential - indeed almost perceptual in nature (Recanati, 2002, 2003) – and which can explain levels (1) to (3); and there is broad consensus that implicatures are recovered inferentially (Huang, 2007). Nevertheless, there is no consensus on the nature of the cognitive processes underlying pragmatics. Relevance Theory (Sperber & Wilson, 1986/1995), to name but one influential school, holds that all pragmatic processes are contextualized inferences that are derived using a single mechanism. The debate in the theoretical and experimental literature is current (e.g. Levinson, 2000; Noveck & Posada, 2003; Bott & Noveck, 2004; Breheny, Katsos & Williams, 2006). There is in contrast broad consensus that whatever is an implicature is recovered by inference (Huang, 2007) and that at least some implicatures are uncontroversially contextualized (i.e. particularized conversational implicatures). Since the debate on the nature of much pragmatic processing is so current, it is politic for a project such as this one to focus on a phenomenon about which there is consensus: a phenomenon such as the particularized conversational implicature.

Experiments and the particularized conversational implicature

As we saw in the General Introduction, suggestive research has been done on implicature in Autistic Disorder (Pijnacker et al, 2009). Using published DSM criteria (APA, 2000) Pijnacker et al investigated how people with autism, people with Asperger's, and typical controls responded to scalar implicatures, finding a

suggestive difference between people with autism and people with Asperger's but no reliable difference between the autistic-spectrum group as a whole and typical controls. The nature of scalar implicatures is hotly disputed, however, with some authors suggesting that scalar implicatures are pragmatic and contextualized (e.g. Sperber & Wilson, 1986/1995), some suggesting that they are pragmatic and generalized (e.g. Levinson, 2000), and still others suggesting that they are at least partly lexicalized hence semantic (e.g. Chierchia, Fox & Spector, *forthcoming*). There are good grounds then to focus on uncontroversially particularized implicatures.

Because few experimental studies have treated particularized conversational implicatures, there is no rich resource of tried-and-tested methods. One method, which approaches the phenomenon obliquely, is that of Surian, Baron-Cohen and Van der Lely (1996). Although this study was sketched in the General Introduction it will prove useful to recapitulate – and expand on – the description.

Surian et al sought to investigate pragmatics in autism by probing how various groups of children respond to violations of Gricean maxims, focusing on the maxims of politeness, relation and quantity. Grice (1969/1989) outlined a general theory of speaker meaning, of which the maxims form an important part. Grice's definition of speaker meaning will be discussed later; here I sketch his account of implicatures, ignoring so-called generalized conversational implicatures and conventional implicatures for the reasons developed above. Grice argued that people's conversations tend to be cooperative, a fact he captured in his Cooperative Principle:

> 'Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged'

(Grice, 1975:45)

Grice offered more specific maxims of conversation, which he believed speakers cleave to in order to behave cooperatively:

Quantity:	1. Make your contribution as informative as required (for the current purposes of the exchange) 2. Do not make your contribution more informative than required.
Quality:	Super-maxim, 'Try to make your contribution one that is true' 1. Do not say what you believe to be false 2. Do not say that for which you lack adequate evidence
Relation:	Be relevant
Manner:	Super-maxim, 'Be perspicuous' 1. Avoid obscurity of expression 2. Avoid ambiguity 3. Be brief (avoid unnecessary prolixity) 4. Be orderly

(Grice, 1975:45-46)

It is worth being as clear as possible about the ways in which Grice believed that speakers can fail to respect these maxims; we will see shortly that Cummings (2007), for one, makes much of these distinctions in critiquing experimental pragmatics and Surian et al in particular. Grice believed that speakers can fail to respect maxims in various ways. Speakers can violate a maxim 'quietly and unostentatiously', which is to say that they can fail to respect a maxim without indicating whether they intended to do so and, if so, what their motivation was. In this case, the failure to respect a particular maxim is arguably linked to a more fundamental failure to respect the maxim of cooperation. Furthermore, maxim violations seem to be extra-linguistic and non-pragmatic, because there is no intention implicitly to communicate the lack of motivation or unwillingness to the hearer. If speakers violate maxims, then they thereby risk breakdowns in

comprehension. In this thesis we will restrict the term ‘violation’ to this specific sense, and not use it more loosely as is sometimes done in the literature. As well as violating a maxim, speakers can opt out of maxims and cooperative principle alike and explicitly declare a lack of will, or perhaps an inability, to continue. They might say, for example, ‘My lips are sealed’. Speakers can also encounter a clash between maxims. For example, a speaker might be unable to give enough information about a subject without also giving information that is poorly evidenced; accordingly the speaker will have to ignore one maxim or the other. And finally – and most importantly with respect to implicature - speakers can flout a maxim, which is to say they can intentionally, and more or less ostensibly, fail to respect a maxim to produce some specific effect.

To test this Gricean framework, Surian et al structured their experiment in the following way. Each experimental trial comprised a question and a pair of responses; participants had to label one response as ‘silly’. One of each pair of responses violated a Gricean maxim; this was the response that Surian et al considered the silly response. The following item from Surian et al’s materials illustrates the design:

3(a)
Question: How would you like your tea?
Answer 1: With milk
Answer 2: In a cup

The first answer ‘with milk’ is straightforwardly acceptable; the second, ‘In a cup’, violates the maxim of quantity and is therefore the so-called silly response. There were also control items, which were ungrammatical: the word order was scrambled.

The Test for the Reception of Grammar (or 'TROG'; Bishop, 1983), a test of verbal comprehension and a standard false-belief task were also administered.

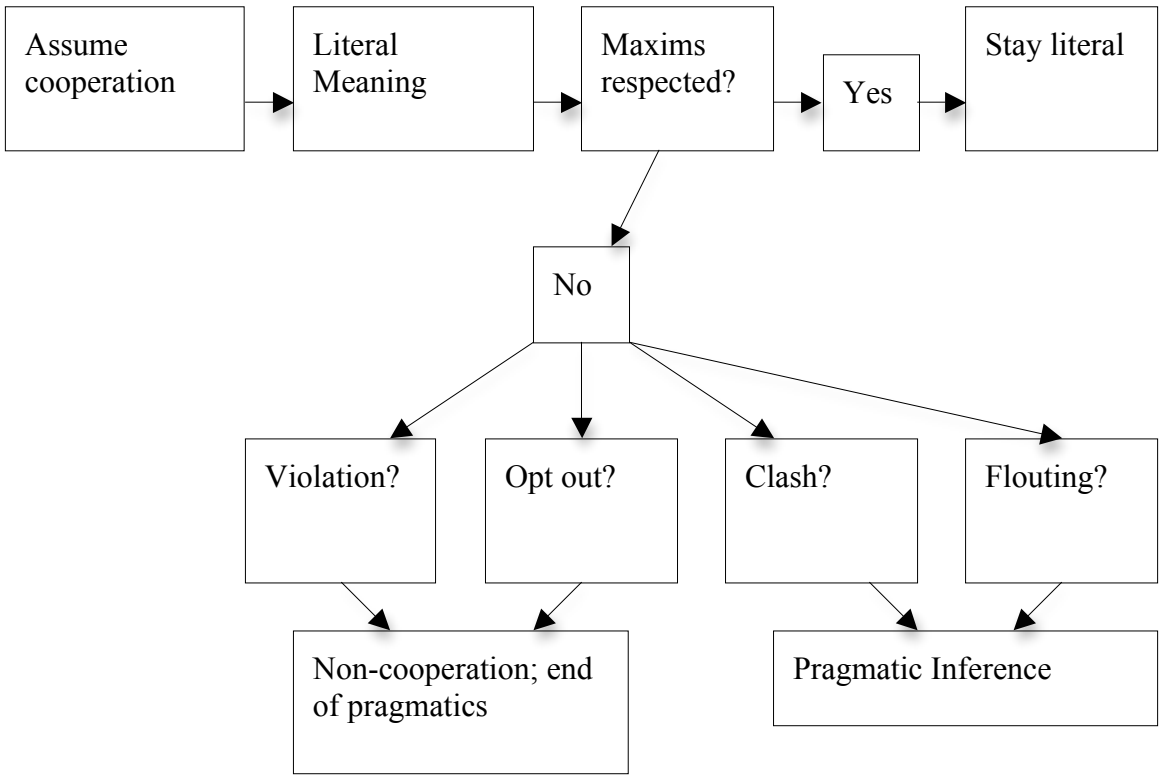
Surian et al ran their experiment with typically developing children, children with specific language impairment, and children with autism; their data are certainly intriguing. Typically developing children, children with SLI and three children with autism, who passed the false-belief test, performed above chance on violations task. The remaining five children with autism performed at chance. The apparent autistic deficit and the association with theory of mind are suggestive.

Suggestiveness aside, Surian et al made some curious decisions about their study, rendering the pragmatic nature of their task, and its fidelity to Gricean theory, questionable. The central concern is this: Surian et al fixed on maxim violation, and yet violation itself may not be pragmatic. Why Surian et al would choose to study maxim violation is unclear, when violation could very well be caused by extra-linguistic motivational problems; interpreting a violation will explicitly call, then, on theory of mind. How important theory of mind is for pragmatics will be discussed below. But even if theory of mind is essential for both violation and flouting, then violation may be harder to interpret because of a paucity of evidence; in flouting there is at least some evidence with which a speaker might postulate a hearer's mental state. In any case Surian et al arguably conflate violation and flouting in many of their materials. Examples like the response, 'In a cup' hardly count as quiet and unostentatious, properties that Grice would seem to require of maxim violations. Rather such examples seem much closer to maxim flouting. Flouting is a better candidate for being pragmatics, but flouting is only carried out

en route to some specific pragmatic goal such as an implicature. Surian et al did not mention pragmatic goals such as implicatures and certainly did not test whether they were produced. But perhaps the most charitable interpretation is that Surian et al saw violation as a basic pragmatic mechanism, that is, as an element that is shared between maxim violation and maxim flouting. However, such a reading does not necessarily follow from Grice (1975).

The contentiousness of Surian et al's assumptions can best be seen in relation to the following figure, based on Cummings (2007).

Figure 3.1: A model of Gricean pragmatics, based on Cummings (2007)



On this, Cummings' (2007) persuasive reading of Grice, the distinctness of maxim violation and maxim flouting is thrown into stark relief. Maxim violation is a separate phenomenon; its occurrence has markedly different implications from that

of maxim flouting. In promoting this reading, Cummings (2007) criticises Surian et al (2007), arguing that they, '[miss] the point of pragmatic exchanges' by focusing on violation and short-circuiting pragmatics. But the significant difference that Surian et al found between typical children and children with ASD is suggestive and well worth exploring further. Surian et al's task can usefully be modified, taking heed of Cummings' strictures, such that the materials allowed participants to derive implicata. So, for example, if participants heard a character ask, 'How would you like your tea?' and another reply, 'In a cup', and if they derived an implicatum such as, 'The second character doesn't want his tea in a mug', then these participants would have engaged in pragmatics.

This section has outlined various doubts about Surian et al's interpretation, that theory-of-mind dysfunction explained pragmatic problems – doubts, indeed, about whether the authors accessed pragmatics at all. The next section will raise some more fundamental theoretical concerns. These concerns originate in Surian et al's claim, not only that a theory-of-mind account does explain the presumed pragmatic deficit, but that only a theory-of-mind account could explain such a deficit. It is not absolutely clear, indeed, that theory of mind plays as definitive a role in pragmatics as authors such as Surian et al suppose. Because this theoretical problem motivates some adaptations made in my experiments I will describe the theoretical problem in some detail.

Theory of Mind and Pragmatics

Surian et al (1996) are not alone in the belief that theory of mind can explain pragmatics; numerous other researchers make a similar case (e.g. Baron-Cohen, 1988; Eisenmajer & Prior, 1991; Happé, 1993), and the argument has philosophical precedent in Grice's definition of speaker meaning:

> 'U meant something by uttering x' is true iff, for some audience A, U uttered X intending:
> (i) A to produce a particular response r
> (ii) A to think (recognize) that U intends (i)
> (iii) A to fulfill (i) on the basis of his fulfillment of (ii)
>
> (Grice, 1969/1989:92; cited in Breheny, 2006)

On Grice's definition, speakers need advanced theory of mind to communicate their meaning effectively (henceforth the term 'communication' is short-hand for communication of the speaker's meaning). Definiens (ii) requires the speaker to reason that, 'If I utter x then A will think that I intend response r'; in other words the speaker must interpret the hearer's belief about his/her (the speaker's) belief. This ability requires second-order meta-representation. But Grice's definition, as various researchers have pointed out, is too broad (e.g. Schiffer, 1972; Strawson, 1964) because it admits non-communicative events such as covert manipulation. Strawson (1964), for instance, gave the following scenario. Person S wants person A to believe P and he tries to bring this about through covert manipulation. S arranges evidence for P, knowing that A is watching him, but also knowing that A does not know that S knows A is watching him. A will not accept the arranged evidence as evidence of P, but will recognise S's intention for A to believe B. Because A has general grounds for assuming that S would not want him to believe P without good reason, A has good reason to accept belief P. Strawson argues that this scenario would count as communication on Grice's definition but should not do so.

Such critics have tightened Grice's definition by adding that the speaker's intentions must be part of mutual knowledge. But in its attitude towards mutual knowledge the literature diverges. Clark (1996), for instance, cleaves to the mutual-knowledge constraint. Sperber and Wilson (1986/1995), in contrast, argue that there are too many obstacles to establishing whether the relevant knowledge truly is mutual, so an alternative constraint is required: that intentions be part of mutually manifest knowledge, where manifestness means that the knowledge is perceptible or inferable. Each of these developments of Grice preserves a commitment to meta-representation; Gricean and post-Gricean pragmatics remain, then, firmly tied to theory of mind.

There are grounds to consider this reliance on theory of mind problematic. For, children who have not yet acquired sophisticated theory of mind, and some adults with autism who never acquire sophisticated theory of mind, nevertheless master basic communication (Breheny, 2006; Gregoromichelaki et al, *forthcoming*). And yet even basic communication, as Breheny notes (*op. cit.*), appears on many pragmatic accounts to presuppose propositional attitudes and sophisticated folk-psychological reasoning. The obvious responses to this problem are (1) to deny that such speakers can really communicate or (2) to assert that these speakers do, in fact, have a tacit theory of mind. But against these responses stand the following points: there seems to be too little evidence to override the appearance of basic communication, and to be too much evidence that children's theory of mind is profoundly different from adults' (Breheny, *op. cit.*). Both obvious responses therefore look untenable.

Breheny (*op. cit.*) makes an alternative response. He posits what Gregoromichelaki et al (*op. cit.*) have called a naïve-optimism strategy: children just assume that interpretations that they derive in context reflect the speaker's intention. A full theory of mind is later grafted onto naïve optimism. But the naïve optimism account is, according to Gregoromichelaki et al (*op. cit.*), problematic because it introduces a two-fold discontinuity in language acquisition. Firstly, by Breheny's own assumption, adult language, the stimulus for language acquisition, is fundamentally different from child language because only adult language is truly communicative in the sense that it properly communicates propositional attitudes; secondly, again by Breheny's own assumption, the posited interpretive mechanism for child language is substantially different from that for adult language; so that it is not at all clear how language acquisition can proceed (Gregoromichelaki et al, *op. cit.*; Kempson, *personal communication*). For the naïve optimism account to hold out against these problems would require a dubious commitment to innateness (Kempson, *personal communication*).

A striking alternative is to deny theory of mind a causative role in communication. A theory that embodies this approach is Dynamic Syntax (Kempson, Meyer-Viol & Gabbay, 2001). Dynamic Syntax uses a form of modal logic to model language comprehension and production in context (*ibid*). The formal details of the theory are irrelevant for present purposes; what is relevant is the evidence of incrementality and emergent mind reading that dynamic-syntax theorists seek to explain. A valuable source of such evidence is split utterances. As Gregoromichelaki et al (*op. cit.*) show, participants in a dialogue readily surrender their utterances to an interlocutor or complete their interlocutor's utterance for them. The following

examples are all due to Gregoromichelaki et al (*op. cit.*). In collaborative completions one speaker anticipates another's meaning and completes his/her utterance for him/her, as in the following examples:

3(b) A: We're going to…
B: Bristol, where Jo lives.

3(c) A: Are you left or…
B: Right-handed?

So far, although the propositions expressed are built up incrementally, they are compatible with Speaker A's intentions, and Speaker B could be consulting Speaker A's intentions in making his/her utterance. But continuations can also be hostile, in which case a new speaker hijacks the original speaker's utterance and completes it with a meaning that might be diametrically opposed to the original speaker's intention, as in the following example:

3(d) A: In fact what this shows is…
B: That you [speaker A] are an idiot.

Here, incrementality and a lack of concern for the original speaker's intentions go hand in hand. Completions can occur at any point in an utterance, even before the speaker's intention could be derived, as in the following example:

3(e) A: They X-rayed me, and took a urine sample, took a blood sample. Er, the doctor…
B: Chorlton?
A: Chorlton, mhmm, he examined me.

In this example, speaker A could have gone on to make any statement about the doctor (e.g. that he took A's blood pressure, that he had ginger hair or supported the England cricket team). Speaker B interrupts to add further information, the doctor's name, and create a fuller proposition. But B's information need not have been part of a proposition that speaker A intended to communicate. Indeed, A may not even

have known the doctor's name; that is, because of the intentional context, B could not assume that A believed both, 'The doctor examined me' and, 'Chorlton examined me'. The success of hostile continuations and completions such as 3(e) is suggestive evidence against a comprehension that relies on clearly defined and recoverable speaker intentions and, in this sense, against a full explanation from theory of mind.

To support this naturalistic evidence, experimental evidence also suggests that speaker's intentions can be emergent from, and hence not causative of, incremental communication. The following experiments are due to Healey (1997) and Mills (2007) and are cited in Gregoromichelaki et al (*op. cit.*). The experiments comprise a maze game using a chat-tool methodology. Participants, who have the same basic information available to them, collaborate in the game by communicating with each other through an instant-chat program akin to MSN Messenger. Participants' messages are routed through the experimenter, who makes occasional interventions, such as a request for clarification, in the guise of one participant or the other. Only the participant to whom the intervention is made sees the intervention; only that participant and the experimenter see the response to the intervention. With this method a natural dialogue can be produced and controlled manipulations introduced to probe participants' intentions. By the end of the task the participants tend to evolve elliptical exchanges whose interpretation depends on the foregoing dialogue; in other words the participants refine the dialogue incrementally. Interventions show that participants' intentions emerge gradually: in early stages the participants clearly communicate but show little evidence of plans or intentions, beyond a vague

intention to collaborate; accordingly, participants avoid attendant negotiation, and if they do engage in such negotiation their performance suffers. Only late in the game can participants' behaviour be consistent with underlying intentions, and even then the elliptical exchanges that were developed incrementally during the task arguably obviate the need to grasp the other participants' intentions. There is thus suggestive experimental evidence that incremental communication can proceed without reading intentions.

The foregoing discussion suggests that the debate on the role of theory of mind in communication is inconclusive: the theory-of-mind and no-theory-of-mind camps can both marshal evidence in support of their views (e.g. Tomasello, 2008; Healey, 1997; Mills, 2007). There are grounds, then, to limit the confidence that we place in attempts, like Surian et al's, to explain pragmatics solely in terms of theory of mind. To be sure, there is a rich literature highlighting the correlation between theory-of-mind dysfunction and pragmatic dysfunction (e.g. Baron-Cohen, 1988; Eisenmajer & Prior, 1991; Happé, 1993;); and it would be hard to argue that theory of mind would not be useful in interpreting utterances. But if those who believe that communication does not entail theory of mind are right (e.g. Gregoromichelaki et al, 2009) then the correlation is hard to interpret. On this view, because theory of mind, at least some of the time, emerges from communication, theory-of-mind dysfunction cannot explain pragmatic dysfunction. Indeed language, pragmatics included, might be one source of evidence for a theory of mind, to complement abilities such as eye-gaze detection that have been credited with providing primitive information for theory of mind to act upon (Baron-Cohen, 1995). There are grounds then to doubt the strength of a causal chain from pragmatic dysfunction to theory-

of-mind dysfunction. Perhaps pragmatic dysfunction forces speakers to rely more on theory of mind, and if, as in autistic spectrum disorders, theory of mind is also dysfunctional, then pragmatic dysfunction and theory-of-mind dysfunction will correlate but be caused independently. The point of this discussion is not to deny that theory of mind can be important in pragmatics, but to emphasise that theory of mind may only ever partly explain pragmatic dysfunction.

Because the studies reported in this thesis involve participants who are likely to struggle with theory of mind - namely adults with autistic spectrum disorders and typically developing children – the dilemma outlined in the previous paragraph poses interpretive problems. Poor performance on theory-of-mind tasks is expected; so too is a correlation between theory-of-mind dysfunction (or naivety) and pragmatic dysfunction (or naivety). Given how well established the findings of such a correlation are, and how difficult they are to interpret, it would add nothing to the literature to find further evidence of the correlation here. What this study could potentially add is a test of an alternative explanation of pragmatic abilities; the next section outlines an alternative way to explain pragmatics, using models of utterance processing.

Models of Utterance Processing and Implicature

Psycholinguistics offers an array of models of utterance processing which might be put to use in explaining pragmatic abilities without directly invoking theory of mind. These models fall into three types: the autonomous model, the direct-access

model, and the graded-salience model. The models differ in how they respond to the question of when contextual information[16] can intervene in utterance processing, that is, whether bottom-up processing must be complete before top-down processing begins or whether top-down processing can interact with bottom-up processing.

The autonomous model identifies lexical processing with bottom-up processing and contextual processing with top-down processing. Its proponents argue that lexical processing happens first and that only after lexical processing is complete are contextual adjustments made; thus the two types of processing do not interact; contextual processing is post-lexical and more demanding than lexical processing (Giora, 2002; Peleg, Giora & Fein, 2004). Gricean pragmatics, as outlined above, seems to fit best with the autonomous model (Grice, 1975). Possible evidence for the implied distinction between lexical and contextual processing can be seen in the datum that reading times are longer for unfamiliar proverbs when they occur with contexts biasing their figurative meanings than when they occur with contexts biasing their literal meanings (Honeck, Welge & Temple, 1998).

The direct-access model differs from autonomous modellers in the matter of interactivity. Direct-access modellers argue that context pervades all of utterance processing; top-down, contextual processing can intervene decisively very early in utterance processing and lexical processing need not be completed first. It follows

[16] It is worth fleshing out the relationship between pragmatics, as defined in this thesis, and the terminology of cognitive modelling that is newly introduced here. I take lexical processing to be equivalent to semantics and contextual processing to be equivalent to pragmatics. Recall that in this thesis pragmatics is defined as the interpretation of implicit meanings intended by the speaker. I mean the transition between 'pragmatics as interpretation of implicit, intended meanings' and 'pragmatics as contextual processing' to imply that contextual processing is a different way of characterizing pragmatics that specifies the type of information used. That is, I understand contextual processing to mean the use of contextual information to recover implicit, intended meanings.

that interpretations are derived only if they are contextually appropriate and that the interpretation of contextually inferred phenomena, such as metaphor and irony, need take no longer than literal meanings (Giora, 2002; Peleg, Giora & Fein, 2004). Representative of this view are Gibbs (1994, 2002) and Wilson and Sperber (2002). Possible evidence for the implied equivalence in processing ease between lexical and contextual processing can be seen in the datum that reading times are equivalent for familiar proverbs and their literal interpretations (Katz & Ferretti, 2001; Turner & Katz, 1997).

The graded-salience model makes a compromise between the two previous models (e.g. Giora, 1997, 2002; Peleg, Giora & Fein, 2004). Graded-salience modellers accept the distinction between bottom-up lexical processing and top-down contextual processing and, like direct-access modellers, argue that the two types of processing happen simultaneously. Uniquely, graded-salience modellers argue that lexical meanings are ordered with respect to salience, a property determined by conventionality, typicality, frequency, and meaning dominance. Although only encoded meanings can be salient, salient meanings do not have to be literal; for instance the figurative meaning of a phrasal idiom can be salient as long as it is encoded. Encoded meanings are activated by the linguistic input, with salient meanings prioritized over non-salient ones; contextual processing affects utterance processing from the start by predicting meanings but cannot initially intervene in lexical processing and cannot block salient meanings. Thus context will affect utterance processing only later on, and contextually inappropriate salient meanings will still be activated even in strongly unfavourable contexts. However, strong contexts can speed up processing so that encoded and inferred meanings can be

accessed equally quickly. The graded-salience model sits well with evidence for both difference and equivalence between lexical and contextual processing because the model predicts both differences and equivalences as a function of salience. So for example, on the graded-salience account it was because Honeck, Temple and Welge's (1998) proverbs were unfamiliar (hence non-salient) that the authors found a difference in reading-time between lexical and contextual processing, and it was because Katz and Ferretti's (2001) and Turner and Katz's (1997) proverbs were familiar (hence salient) that these authors found equivalence in reading-time between lexical and contextual processing.

This modelling approach is markedly different from the theory-of-mind-driven approach taken by Surian et al (1996): no mention has been made of theory of mind whatsoever in this section. By applying the competing models to typical adults, typically-developing children and adults with ASD, we can potentially gain considerable insight into pragmatics in typical and atypical development. The experiments reported in this thesis are an attempt to apply these models to implicature, the paradigm case of pragmatics. The next section will outline the experimental design.

Experiment Design

The experiments reported in this thesis were designed in light of the criticisms of Surian et al (1996), of the limitations of the theory-of-mind approach, and of the modelling approach. The experiments took the following form.

Particularized conversational implicatures were deployed in a task based on Surian et al (1996), in which participants heard micro-conversations consisting of a question-response pair. Unlike in Surian et al's task participants heard only one response for each question; response types were counterbalanced between participants. Responses varied in the extent to which they involved implicature. Consider the following example:

3(f)

Speaker X: How would you like your tea?
Speaker Y: (i) Without any milk, please *or*
(ii) Preferably in a cup
(iii) With a little lettuce

Response (i) is literal and salient; response (ii) is pragmatic because its appropriateness depends on an implicatum such as, 'Please don't spill the tea everywhere'; response (iii) is a nonsense control. Together these conditions comprised an independent variable, namely Meaning Type, with three levels (nonsense, salient, and pragmatic). There is an important departure from Surian et al's (1996) design in that response (ii), in Gricean terms, a maxim is flouted rather than violated; put in less theoretically loaded terms whether a response is accepted depends on implicature and not on detecting an abnormality.

Recall though that particularized conversational implicatures are contextually mandated, and that models of utterance processing make their distinctive predictions about contextual information. Accordingly a second independent variable, Context Type, was needed. This variable had three levels: null; congruent; and incongruent. While the null-context condition is self-explanatory the other conditions require further explanation. A context was congruent if it agreed with the

response that followed it and incongruent if the meaning type disagreed with the response. Consider the following examples:

3(g) Salient Meaning, Congruent Context Condition
Context: X's writing a shopping list
Speaker X: Now what shall we eat for lunch?
Speaker Y: I think, cheddar cheese and grilled bacon

In this item there is nothing to surprise the participant. There is a context and there is a salient response; context and response agree.

3(h) Pragmatic Meaning, Congruent Context Condition (Pragmatic Context with Pragmatic Response)
Context: The dog loves having walks
Speaker X: Has someone walked the dog?
Speaker Y: He does look very happy

In this item there is, again, nothing surprising: there is a context and a pragmatic response, and context and response agree: taken together, context and response implicate that the dog looks happy, indeed is happy, because he has been taken for a walk.

3(i) Salient Meaning, Incongruent Context Condition

Context: X just dropped her/his cake.
Speaker X: How would you like your chocolate cake?
Speaker Y: With some cream on it

This item is more surprising. There is a context and a salient response, but context and response do not agree: that is, Speaker Y makes no reference to the dropping of the cake.

3(j) Pragmatic Meaning, Incongruent Context Condition (Salient Context with Pragmatic Response)
Context: Speaker Y has very demanding tastes
Speaker X: How would you like your hair cut?
Speaker Y: With a pair of scissors

Once again, a surprising item: context and pragmatic response do not agree. Speaker Y makes a response which is unusual without specific contextual justification. If, for example, the context had said that 'Speaker X usually cuts people's hair with clippers, hence very short', then Speaker Y's response would be pragmatically licensed; here, it is not.

The forced-choice format favoured by Surian et al (1996) was rejected with a view to conducting the experiment with children and participants with ASD because such formats have been known to over-estimate participants' competence. Instead participants rated responses on a Likert-style scale from 1 to 4 for how much sense they made. This scale also allowed subtler comparisons between conditions. One dependent variable then was the participant's rating. Participants' response times were also recorded, forming a second dependent variable.

A confound in the materials?

There is a potential problem with the experimental design that, if proved actual, would cause a confound that would rule out testing the predictions described above. The problem concerns what might be term Contextual Priming and Resonance. These phenomena can be made clearer with the following examples.

3(k)	Context:	Y very much dislikes milk.
	Speaker X:	How would you like your tea?
	Speaker Y:	I'd like it without milk.

3(l)	Context:	X just spilt milk everywhere.
	Speaker X:	How would you like your tea?

Speaker Y: Preferably in a cup, please[17].

Example 3(k) directly primes the word ‘milk’ in the response, and because the context in 3(l) only indirectly primes the word ‘cup’, the design confounds Contextual Priming and Meaning Type. While the criticism does indeed hold for the examples mentioned above, this distinction in priming was not systematic; that is, there were multiple cases of indirect primes for salient meanings and direct primes for pragmatic meanings. The critical question, then, is in what proportion of cases did contextual priming favour semantic meanings? There were 40 questions, and 2 possible contexts. Thus there were 80 pairings of semantic and pragmatic responses. There were 8 potentially problematic pairs: a difference in contextual priming did not systematically co-occur with a difference in meaning type.

The potential for a confound owing to Resonance is also seen in the examples above. In example 3(k) the word ‘like’ occurs in both question and response, but in example 3(l) it occurs only in the question. Thus, the response in example 3(k) is more resonant with the question than is the response in example 3(l). If salient responses were systematically more resonant with questions than pragmatic responses, then the design would confound Resonance and Meaning Type. But as for contextual priming, there are multiple cases of salient responses with words that are at best indirectly primed by questions and of pragmatic responses with words that are directly primed by the question. Again, there were 80 pairings of semantic and pragmatic responses to consider. There were 6 potentially problematic pairs: a

[17] This response depends on irony as well as implicature and may therefore add an extra level of complexity to comprehension; this item is unique amongst the materials, however, in being ironic.

difference in resonance did not systematically co-occur with a difference in meaning type.

Although future studies should replace these problematic items, the present study does not confound Meaning Type with Contextual Priming or Resonance.

Predictions

Predictions from each model are shown in Table 3.1. In some cases, predictions were not straightforward, because of ambiguities in the competing models. These predictions reflect my impression of the most likely reading of the models. It also is worth noting that autonomous and direct-access models only make predictions for salient meanings insofar as they are also literal. However, all the salient meanings used in these experiments are also literal meanings.

Table 3.1: Experimental predictions

=	**Autonomous Model**	**Direct Access**	**Graded Salience**
A. Differences between conditions when no context			
(1) Salient vs. pragmatic	Salient faster and rated higher	Salient faster and rated higher	Salient faster and rated higher
(2) Pragmatic vs. nonsense	Pragmatic slower but rated higher	Pragmatic slower but rated higher	Pragmatic slower but rated higher
B. Differences between salient and pragmatic meanings			
(3) Congruent Context:	Both rated good, but salient higher; salient faster	No difference in reaction time or rating	Salient faster and rated higher

(4) Incongruent context	No difference in reaction time or rating; both slow and bad	No difference in reaction time or rating: both slow and bad	Salient should be rated higher than pragmatic; should both be slow
C. Effects of context on salient meanings			
(5) no context vs. congruent context	Congruent context slower; no change in rating	Congruent context faster and higher rating	Congruent context faster and higher rating
(6) no context vs. incongruent context	Longer reaction times and lower ratings for incongruent contexts	Longer reaction times and lower ratings for incongruent contexts	Longer reaction times but good ratings for incongruent contexts
D. Effects of context on pragmatic meanings			
(7) no context vs. congruent context	Congruent context slower but higher ratings	Congruent faster and higher ratings	Congruent faster and higher ratings
(8) no context vs. incongruent context	Incongruent context slower; ratings lower	RTs equivalent; ratings lower	RTs equivalent; ratings lower

Since many of these predictions are non-transparent it is worth fleshing them out. The first set of predictions, (A), concerns differences between meaning types when no context was present. Here the models make similar predictions. It is intuitively plausible that salient meanings should be processed fastest and rated highest, since all models allow for default meanings. Less clear is how the models would handle pragmatic responses as opposed to nonsense: should pragmatic responses be preferred to nonsense, or should they be considered as bad as nonsense? In Table 1 I assume that participants hypothesize contexts to make pragmatic responses acceptable.

The second set of predictions, (B), concerns the difference in processing salient and pragmatic responses. Recall that salient meanings are also literal meanings: the autonomous and direct-access models made the same predictions for salient

meanings as they would have done for literal meanings in general. For congruent contexts the autonomous and graded-salience models make the same prediction: that salient meanings would be rated faster and higher than pragmatic meanings. The direct-access model predicts no difference, because on that view context is decisive from the outset. For incongruent contexts, the autonomous and direct-access models make the same prediction: that there would be no difference in reaction time or rating between salient and pragmatic meanings, and that both would be rated slowly and as poor. This reading of the autonomous model is disputatious because of an ambiguity in the autonomous model. The autonomous model suggests that interpreters always start with the salient (or any literal) meaning and then make any necessary inferences after detecting contextual mismatch. It is not clear what happens to the literal meaning after this mismatch is detected: is it discarded or kept as a back-up? We assume the simplest option: that the salient (literal) meaning is discarded after a mismatch with context is detected. On this reading, the autonomous model predicts that there should be no difference in rating speed or level, because with either meaning type there will be a contextual mismatch, demanding additional processing to attempt to recover a sensible meaning. The graded-salience model, in contrast, predicts that salient meanings should be activated irrespective of their contextual appropriateness, and therefore that salient meanings should be rated as acceptable. Since the incongruent context should lead participants down the garden path, however, the rating should be slow.

The third set of predictions, (C), treats the effects of context on salient meanings. Consider the effect of adding a congruent context. Here, the autonomous model makes a distinctive prediction: that congruent contexts should increase the rating

time and should not change the rating level. The additional processing time comes from the time taken to process the additional, contextual information; the lack of a difference in ratings comes from the fact that a salient meaning is already maximally acceptable without context such that a congruent context should make no difference to the rating level. In contrast, the graded-salience and direct-access models predicts that context should intervene early in interpretation and facilitate it: hence, that congruent contexts should increase processing speed and rating level. Consider, now, the effect of adding an incongruent context. The autonomous and direct-access models straightforwardly predict that reaction ratings would take longer and would be lower than for the null context condition. The graded-salience model predicts that ratings would take longer but that ratings would remain high for salient meanings in incongruent contexts, because salient meanings are supposed to be activated automatically despite the influence of opposing contexts.

The final set of predictions, (D), treats the effects on pragmatic meanings of adding contexts. For congruent contexts, the autonomous model predicts that ratings should take longer to deliver, because of the additional time to process the context, but that the context should improve ratings of the pragmatic meaning. The direct-access and graded-salience models both predict that adding contexts should decrease the time taken to deliver a rating, because of an early and facilitative effect of context, and that ratings should improve. For incongruent contexts, the autonomous model predict slower and lower ratings; the direct-access and graded-salience models predict no change in reaction time but lower ratings, since the opposing context should rule out the pragmatic reading early on in utterance processing.

Summary

This chapter has outlined an experiment that was designed to test cognitive models of utterance processing in various populations. The experiments differ markedly from Surian et al (1996) in that they directly invoke cognitive models, rather than theory of mind, to explain pragmatic abilities. The justification for taking this modelling approach was that it avoids the problems associated with theory-of-mind explanations that were outlined above. The experimental design improves on Surian et al in that it accesses properly pragmatic processing by allowing implicature. The design therefore had the potential to reveal interesting differences between populations of speakers.

Chapter 4: Study of Typical Adults

Introduction

The adult study was carried out with the following aims: (1) to address the need for an experimental study which employs conversational implicatures that are uncontroversially particularized; (2) to adapt Surian et al's (1996) design so that it more clearly tests pragmatics, by focusing on maxim flouting rather than maxim violation; (3) to test among models of utterance processing both as an end in itself and as a possible alternative to theory-of-mind explanations of pragmatics; (4) to establish a yardstick against which to measure clinical and developmental groups.

Because the experimental design and predictions were thoroughly described in the preceding chapter they will not be reprised here; instead we turn to the methods section.

Methods

Design

The study comprised a single experiment using a mixed design with two dependent variables. The first dependent variable was ratings of acceptability on a scale from 1 to 4; the second, response time. Independent variables were the Meaning Type (Nonsense, Salient, Pragmatic) and Context Type (Null, Congruent, Incongruent).

Participants

26 normal adult speakers of English as a first language gave informed consent and took part in this study. They were principally recruited from the staff of a school and a local community in Rutland; some participants were students at the University of Oxford. Care was taken to include as broad a range of educational backgrounds as possible in a study of this size. Participants had left education at the following levels: GCSE or equivalent (n = 4); A-level or equivalent (n = 5); vocational training (n = 5); undergraduate degree (n = 12). Each participant was randomly assigned to one of two groups depending on which task s/he would complete, (judging as sensible, or as likely to be produced by a native speaker, on a scale from 1 to 4). The total group each comprised 16 women and 10 men. All participants were neurologically normal and free from developmental or acquired language disorders.

Materials

The experimental materials were 40 micro-conversations consisting of question-response pairs and possible contexts. Some of the questions and responses were lifted from Surian et al (1996), others were variations on Surian et al's materials, and still others were developed entirely from scratch. Each question was paired with one of three possible responses: there was a nonsensical response, a salient response, and a pragmatic response. The question-response pair could be presented with or without a context. When a context was present it could be congruent or incongruent with the response: In this way the conversations manipulated two variables: Meaning Type (Nonsense, Salient, Pragmatic) and Context Type (Null, Congruent and Incongruent).

The following example shows all possible responses and contexts for one of the questions:

4(a)		
Speaker X:	Y,	How would you like your tea?
Speaker Y:	(nonsensical)	I'd like it over French fries.
	(salient)	I'd like it without milk.
	(pragmatic)	I'd prefer in a cup.
Preceding context:	(congruent with salient)	Y very much dislikes milk.
	(congruent with pragmatic)	X just spilt tea everywhere.

Some participants saw the above example in the following configuration:

4(b)	
Context:	X just spilt tea everywhere.
Speaker X:	Y, How would you like your tea?
Speaker Y:	I'd prefer in a cup.

In this configuration, speaker Y makes a pragmatically acceptable response which is licensed by the context: the context is congruent with the response. In an alternative configuration which other participants heard, the same question and response pair is preceded by the context, 'Y very much dislikes milk', a context which promotes a salient response: that context is incongruent with the response.

The number of words in all components of the conversation was strictly controlled, and to avoid item-specific effects each question was presented to participants only once and thus appeared in only one configuration. Each configuration was presented to at least one participant in the sample. To avoid order effects, items were counterbalanced using the Latin-square method such that each participant heard the questions in a different order.

Procedure

Experimental sessions occurred in a quiet environment either in the participant's home or in the Department of Experimental Psychology at the University of Oxford. Having already been randomly assigned to one of two groups, participants were asked to give a rating on a scale from 1 to 4 to Speaker Y's response in each micro-conversation. The first group was asked to give ratings according to how much sense Y's response made: if the response made perfect sense, participants were to give a rating of 4; if the response was nonsensical, a rating of 1. The second group was asked to give ratings according to how natural it would be for a native speaker of English to make Y's response: if very natural, participants were to give a rating of 4; if very unnatural, a rating of 1. Since preliminary analyses showed no difference between task types, the data are collapsed in the following sections. Participants were asked to make their decision quickly and intuitively, and to avoid deliberating over it. They were asked to make their response simply by saying the appropriate number. The experimenter then read each micro-conversation aloud, using a stopwatch to time the interval between the end of the last word of the response and the start of the participant's response. A stopwatch was seen as providing sufficient accuracy given that differences between conditions were in the order of half a second rather than milliseconds. Using a stopwatch also minimised the necessary equipment for the later autism study, so that participants could feel as comfortable as possible.

Results

Average ratings were calculated by condition for each participant. The analyses are organized according to the predictions in Table 3.1.

Prediction Set (A): No Context; comparison of three meaning types

The first set of analyses concerned the predictions from section (A) of Table 3.1. Table 4.1 shows descriptive statistics for Meaning Type without context.

Table 4.1: Typical adult data - Mean (SD) ratings and RTs for meaning type without context

		Rating		RT	
	N	Mean	Std Deviation	Mean (s)	Std. Deviation
Nonsense	26	1.24	.40	1.97	1.92
Salient	26	3.79	.24	1.70	1.02
Pragmatic	26	2.92	.56	2.47	1.23

A 3 x 2 ANOVA was performed on these data. Average Rating was the dependent variable, and Meaning Type (nonsense, salient, pragmatic) was the independent variable. The assumptions of sphericity and of equal variance were satisfied for all variables and their interactions. There was a significant main effect of meaning type ($F(1,24) = 361.61$, $p < .05$, partial $\eta^2 = .94$): pairwise comparisons revealed that participants significantly preferred salient meanings to pragmatic meanings ($p < .05$) and pragmatic meanings to nonsense ($p < .05$).

The RT data were entered into a 3 x 2 mixed ANOVA, with Reaction Time as dependent variable and Meaning Type (nonsense, salient, pragmatic) as independent variable. The assumptions of sphericity and of equal variance were satisfied. The main effect of meaning type was significant ($F(1,24) = 4.84$, $p < .05$, partial $\eta^2 = .17$): pairwise comparisons revealed that participants processed salient responses significantly faster than nonsense responses ($p < .05$), which in turn they processed significantly faster than pragmatic responses ($p < .05$).

Prediction Sets (B) and (C) With Context

General Analysis of Ratings Data Figure 4.1 summarizes the descriptive statistics for Meaning Type with context. Because the nonsense condition above revealed a floor effect, that condition was removed from further analysis in this section.

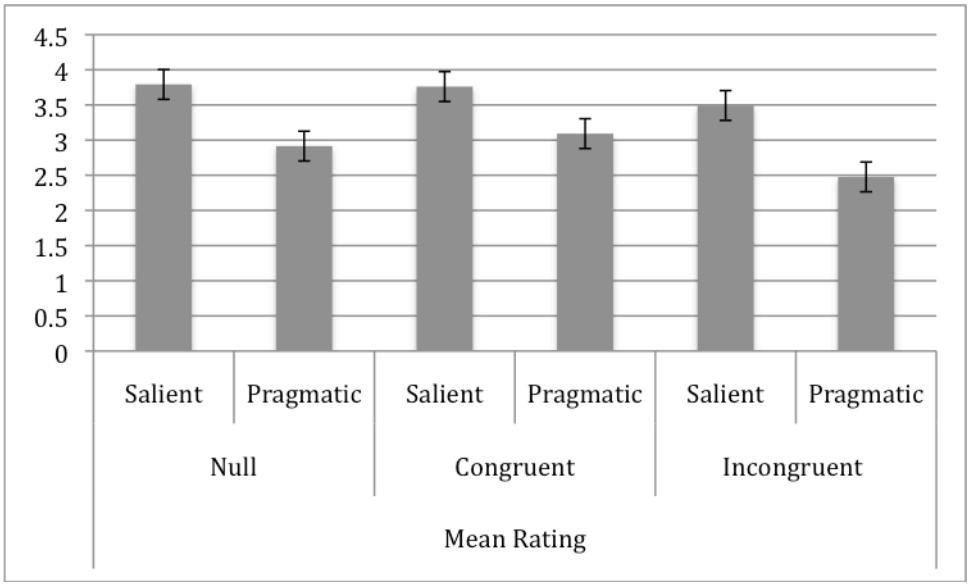


Figure 4.1: Typical adult data - Mean rating by Context Type (Null, Congruent, Incongruent) and Meaning Type (Salient, Pragmatic); error bars are standard error

A 2 x 3 ANOVA was performed on these data, with average rating as the dependent variable and Meaning Type (salient, pragmatic) and Context Type (none, congruent, incongruent) as independent variables. The assumption of sphericity was satisfied for all variables and their interactions, as was the assumption of equal variances. There was a significant main effect of Meaning Type ($F(1,24) = 106.88$, $p < .05$, partial $\eta^2 = .82$): participants' ratings were significantly higher for the Salient condition ($M = 3.68$, $SE = .05$) than for the Pragmatic condition ($M = 2.83$, $SE = .10$). There was also a significant main effect of Context Type ($F(2,25) = 22.54$, $p < .05$, partial $\eta^2 = .48$): pairwise comparisons with the Bonferroni adjustment revealed that participants' ratings were highest for the Congruent Context condition

(*M* = 3.43, *SE* =.08) but not significantly different from the Without Context condition (*M* = 3.35, *SE* = .07), and that ratings for both these conditions were significantly higher than those for the Incongruent Context condition (*M* = 2.99, *SE* = .09). There was also a significant interaction between Meaning Type and Context Type (*F* (2,48) = 3.71, *p* <.05, partial η^2 =.13).

General Analysis of Reaction Times Figure 4.2 summarizes the descriptive statistics for reaction-time data for meaning type with context.

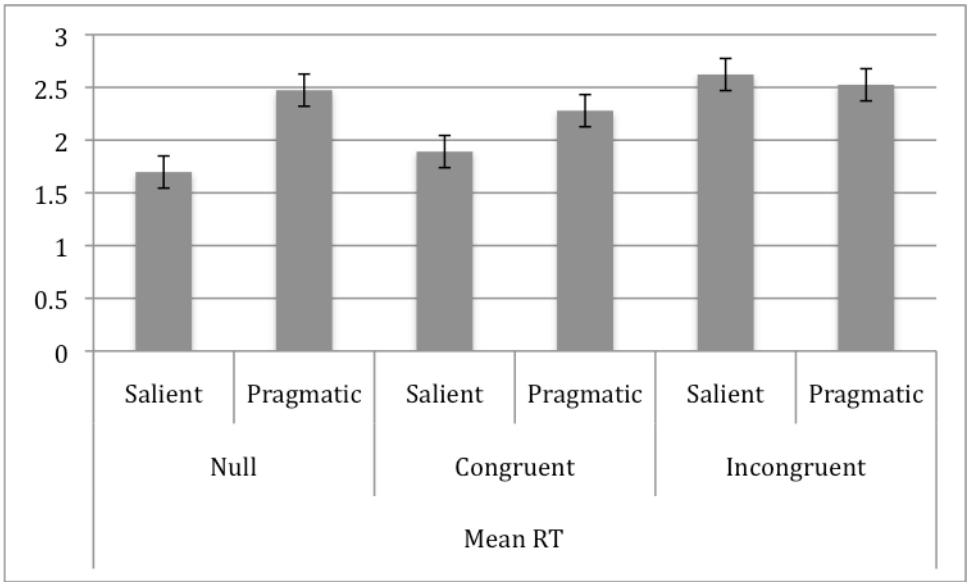


Figure 4.2: Typical adult data - Mean RT (s) by Context Type (Null, Congruent, Incongruent) and Meaning Type (Salient, Pragmatic); error bars are Standard Error

These data were entered into a 2 x 3 x 2 mixed ANOVA, with Reaction Time as dependent variable, Meaning Type (salient, pragmatic) and Context Type (null, congruent, incongruent) as within-subjects variables, and task type (sense, foreign speaker) as between-subjects variable. The assumptions of sphericity and of equal variance being satisfied, the analysis yielded the following results. There was a significant main effect of Meaning Type (*F*(1,24) = 4.84, *p* <.05, partial η^2 =.17): the Salient condition produced significantly faster reaction times (*M* = 2.07s, *SE* = .26) than the Pragmatic condition (*M* = 2.43, *SE* = .22). There was also a significant

effect of Context Type (F (2,24) = 9.49, p <.05, partial η^2 =.28): the Null Context (M = 2.09, SE = .21) and Congruent Context (M = 2.09, SE =.24) conditions were not different from each other, but the Incongruent condition produced slower reaction times (M = 2.57, SE =.28). The interaction of Meaning Type and Context Type was significant (F(2,48) = 9.49, p <.05, partial η^2 =.18). Figure 3, below, represents the interaction.

To investigate the predictions more closely, paired-sample t-tests were performed; because all predictions were directional, one-tailed tests were used.

Prediction Set (B)

In Congruent contexts, Salient meanings were rated as significantly better than Pragmatic meanings in Congruent contexts (t(25) = 8.19, p <.05); and salient meanings were rated significantly more quickly than Pragmatic meanings were (t(25) = 2.29, p <.05). In Incongruent contexts, Salient meanings were rated as significantly better than Pragmatic meanings (t(25) = 9.00, p <.05); and there was no significant difference in reaction time (t(25) =.33, p =.75).

Prediction Set (C)

When Congruent Contexts were added to Salient meanings, there was no significant difference in rating (t(25) = .51, p = .31), nor was there a difference in reaction time (t(25) =1.61, p =.51). On the other hand, when Incongruent contexts were added to Salient meanings, there was a significant decrease in rating (t(25) = 4.62, p <.05) and a significant increase in reaction time (t(25) =3.94, p <.05).

Prediction Set (D)

When Congruent contexts were added to pragmatic meanings, there was a marginally significant increase in rating ($t(25) = 1.55$, $p = .07$) and no significant difference in reaction time ($t(25) = 1.20$, $p = .12$). When Incongruent contexts were added to Pragmatic meanings, there was a significant decrease in rating ($t(25) = 3.61$, $p < .05$) and no significant difference in reaction time ($t(25) = .38$, $p = .36$).

Discussion

This study with typical adults met a variety of aims that were outlined in the preceding chapter. The study offered a rare glimpse into the particularized conversational implicature in an experiment, and it adapted Surian et al's design by focusing on maxim flouting and by allowing participants to respond with a Likert-style scale. The satisfaction of these aims was covered in some detail in the preceding chapter. This section discusses how the remaining aims were satisfied: first it treats the question of how the data bear on the models of utterance processing (, and as an aside how the evidence for the models bears on leading linguistic theories of pragmatics); secondly it discusses how reliable the data are as a point of comparison with clinical and developmental groups.

The following discussion will track the predictions as they were presented in the Table of Predictions on page 93/4.

Prediction Set (A)

This set comprised theory-neutral predictions for Meaning Type without context. Participants gave salient meanings the highest ratings, followed by pragmatic meanings; they gave nonsense low ratings. The pattern was robust, with a large effect size. Participants rated salient meanings fastest, then nonsense, and then pragmatic meanings. These data support predictions (1) and (2): they suggest that participants hypothesized contexts to make pragmatic responses more acceptable than nonsense; recall that the pragmatic responses were pragmatic in virtue of particularized conversational implicatures, which depend on a context. Since participants might have equated pragmatic responses with nonsense, this is an interesting finding: it suggests insofar as this task reflects natural communication that interpreters are tolerant of pragmatic responses even without a supporting context, preferring to commit to the extra work of deriving a licensing context than to write such responses off as nonsense. The reaction-time data suggest that interpreters first suppose nonsense to be non-salient or pragmatic meaning: they look for possible licensing contexts and, when they fail to identify such contexts, they finally reject the responses as nonsensical. When licensing contexts can be found in contrast the additional processing time is needed to deploy the context as a premise and use it to derive the actual interpretation. This process is compounded by the likelihood of multiple contexts for a pragmatic response, demanding that most appropriate context be selected.

Prediction Set (B)

These predictions compared the effects of context on salient meanings with those on pragmatic meanings. Prediction (3) separated the direct-access model from the autonomous and graded-salience models: the former predicted that, in congruent contexts, there should be no differences in processing-time or rating for salient and pragmatic meanings. In fact there was a significant difference, with salient meanings rated significantly higher ($t = 8.19$, $p < .05$) and significantly faster ($t = 2.29$, $p < .05$) than pragmatic meanings. This first piece of evidence, then, does not support direct access. Notably, though, one could save direct access by arguing that contexts were insufficiently strong to equalize processing-times of salient and pragmatic responses. As this point shows, the choice between models will have to be an inference to the best explanation, based on how many of the predictions are straightforwardly satisfied by each model.

Prediction (4) distinguished between the graded-salience model, on the one hand, and the autonomous and direct-access models on the other. The prediction concerned how salient and pragmatic meanings were rated in incongruent contexts. Because the graded-salience model holds that salient but inappropriate meanings will be automatically activated by the input, the model predicted that salient meanings should be rated higher than pragmatic meanings. Reaction times should both be slow. The autonomous and direct-access models predict that salient and pragmatic meanings should be rated as equally bad, equally slowly. The data from this study support the graded-salience model: while both types of meaning were

slow to process, salient meanings were rated as significantly better than pragmatic ones (t =9.00, p <.05).

Prediction Set (C)

These predictions treated the effects of context on salient meanings. Prediction (5) separated the autonomous model from the direct-access and graded-salience models: the autonomous model predicted that adding a congruent context to salient meanings should slow down reaction times and should not change rating; the direct-access and graded-salience models predicted in contrast that adding a congruent context should speed up reaction times and increase ratings. The data are inconclusive for this prediction. On the one hand there was no significant difference in rating, supporting the autonomous model. On the other hand there was no significant difference in reaction time, suggesting a real-time increase in processing speed and therefore supporting direct access and graded salience. It is quite plausible, as suggested for prediction (3), that the contexts were insufficiently strong to equalize ratings. Alternatively, there may have been a ceiling effect: salient meanings were rated so high, that is, that a significant increase was impossible. These data are therefore taken to support direct access and graded salience.

Prediction (6) teased the graded-salience model apart from the other models: graded salience predicts that salient meanings in incongruent contexts should have longer reaction times but good ratings; the other models both predicted longer reaction times and low ratings. Incongruent contexts did indeed decrease the acceptability of salient meanings (t = 3.61, p <.05) and increase reaction times (t = 3.95, p <.05);

but, crucially, ratings remained high for salient meanings in incongruent contexts, as predicted by the graded-salience model. Salient meanings seem to come for free, in a sense, despite opposing contexts.

Predictions Set (D)

These predictions treated the effects of context on pragmatic meanings. Prediction (7) compared how pragmatic meanings were rated in null contexts with how they were rated in congruent contexts. Here, the autonomous model predicted that congruent contexts should increase ratings and reaction times, whereas the direct-access and graded-salience models predicted that congruent contexts should increase ratings and decrease reaction times. The data support the direct-access and graded-salience view. Ratings were indeed higher for pragmatic meanings in congruent contexts, as shown by the marginally significant trend ($t = 1.55$, $p = .07$). There was no significant difference in processing-time between null and congruent contexts; however, this lack of a statistical difference suggests a real-world difference. Other things being equal the contexts should slow down reaction times because they add information; if there is no significant difference, this suggests that the additional information can speed up processing to such an extent that the effect of processing the context is negated. There is no obvious way for the autonomous model to account for these facts.

Prediction (8) parts the autonomous model from the direct-access and graded-salience models. If the autonomous model were true, then pragmatic meanings should take longer to process when in incongruent contexts than when alone. But that is not the case: in fact, participants were as fast to deliver their low ratings of

pragmatic meanings in incongruent contexts as they were without contexts. This fact makes sense if contextual information can begin early and can facilitate processing, and this is precisely what both direct-access and graded-salience models allow for. The data are therefore taken to support those models.

Modelling Implications

The data outlined above do not unequivocally evidence any one model. Nevertheless, a distinction did emerge in favour of the graded-salience model. The data for prediction 3 support the autonomous and graded-salience models; for prediction 4, the graded-salience model; for prediction 5, the direct-access and graded-salience models; for prediction 6, the graded-salience model; and, finally, for prediction 7 and 8, the direct-access and graded-salience models. On balance, then, the graded-salience model of utterance interpretation fares best at capturing the current data, there is also some good evidence for the direct-access model, and there is poor evidence for the autonomous model.

Table 4.2: Predictions met

=	**Autonomous Model**	**Direct Access**	**Graded Salience**
A. Differences between conditions when no context			
(1) Salient vs. pragmatic	Salient faster and rated higher ☑	Salient faster and rated higher ☑	Salient faster and rated higher ☑
(2) Pragmatic vs. nonsense	Pragmatic slower but rated higher ☑	Pragmatic slower but rated higher ☑	Pragmatic slower but rated higher ☑
B. Differences between salient and pragmatic meanings			
(3) Congruent Context:	Both rated good, but salient higher; salient faster ☑	No difference in reaction time or rating ☒	Salient faster and rated higher ☑
(4) Incongruent context	No difference in reaction time or rating; both slow	No difference in reaction time or rating: both slow	Salient should be rated higher than pragmatic; should

	and bad ☒	and bad ☒	both be slow ☑
C. Effects of context on salient meanings			
(5) no context vs. congruent context	Congruent context slower; no change in rating ☒	Congruent context faster and higher rating ☑	Congruent context faster and higher rating ☑
(6) no context vs. incongruent context	Longer reaction times and lower ratings for incongruent contexts ☒	Longer reaction times and lower ratings for incongruent contexts ☒	Longer reaction times but good ratings for incongruent contexts ☑
D. Effects of context on pragmatic meanings			
(7) Pragmatic: no context vs. congruent context	Congruent context slower but higher ratings ☒	Congruent faster and higher ratings ☑	Congruent faster and higher ratings ☑
(8) no context vs. incongruent context	Incongruent context slower; ratings lower ☒	RTs equivalent; ratings lower ☑	RTs equivalent; ratings lower ☑

Theoretical Implications

The data outlined here bear not only on models of utterance processing but also, by extension, on pragmatic theory. Just how these data might influence theory can be illustrated with respect to Relevance Theory. Relevance Theory was discussed in a preceding chapter; this section will briefly reprise the fundamentals of the theory. Relevance-theoretic pragmatics falls out from the view that human cognition is geared towards a search for relevance, where relevance is a function of maximised cognitive effects, such as a true conclusion that could only be drawn from input and context together (Wilson & Sperber, 2002), and minimised processing effort. Relevance comes in degrees; accordingly hearers follow a path of least effort in

interpreting an utterance, fixing on the first interpretation that satisfies their expectation of relevance (*ibid*).

While Relevance Theory recognizes a distinction between decoding meaning and inferring it, the distinction does not imply that decoding and inference happen consecutively. Relevance Theory suggests that hypotheses are derived, favoured and rejected throughout interpretation: to quote, 'there is no suggestion that the literal meaning [of an utterance] must be tested first' (Wilson & Sperber, 2002:63). It is presumably at least in part consideration of direct-access data that has led relevance theorists to posit a kind of massive interactivity whereby context, literal content and cognitive effects (including implicated premises and conclusions) are mutually adjusted in any order (*ibid*). Relevance Theory, then, seems wedded to a direct-access model of utterance processing.

To the extent that Relevance Theory implies a direct-access model and this study implies a graded-salience model, the two are in conflict. This study suggests, albeit somewhat weakly, that Giora and Fein (1999a, b) and Honeck, Wedge and Temple (1998) are right in arguing that the direct-access model of utterance processing does not capture important data. This study instead supports, again somewhat weakly, a model in which there is a privileged role for lexical meanings, structured around salience, and in which contextual inference cannot initially intrude on lexical processing. Importantly, though, this study in no way contradicts the core tenets of Relevance Theory: it does not, for instance, speak against relevance principles, findings in Lexical Pragmatics or the experimental evidence offered in support of Relevance Theory (e.g. Van der Henst & Sperber, 2004). Given the compatibility of

a graded-salience model with Relevance Theory, and in spite of Relevance theorists' favour towards direct-access models, it is unclear how committed Relevance Theory must be to a direct-access model.

It is instructive to consider how Relevance Theory might look if underpinned by a graded-salience model. Graded-salience Relevance would still have place for a code, for pragmatically enriched explicit meanings (explicatures), and for pragmatically derived implicatures. Furthermore, pragmatic processes such as lexical narrowing and broadening could still intervene in lexical access. The changes to Relevance Theory would focus on details of how and when such elements interact. Graded salience would demand an early primacy of salient encoded meanings and a phase of impenetrability to contextual inference. *Contra* Wilson and Sperber (2002) then, there would be a suggestion that salient meanings (hence many literal meanings) are given initial priority in the online generation of meaning hypotheses. Such changes would probably have limited effect on the intuitive appeal of Relevance Theory, or even on the application of Relevance Theory to naturalistic linguistic data. But since Relevance Theory is intended to be a cognitive-psychological theory of meaning (Sperber & Wilson, 1986, 1995), the precise details of modelling – of the implementation of the theories – are of substantial theoretical importance.

Clinical and Developmental Implications

A separate aim of this study was to establish points of comparison for studies with clinical and developmental groups. The following key points arise from the discussion above for the ratings data:

(1) Without context, salient meanings are rated best, followed by pragmatic meanings, followed by nonsense
(2) Congruent contexts maintained or improved salient meanings
(3) Incongruent contexts lowered ratings for salient meanings, but these ratings were still high
(4) Congruent contexts improved ratings for pragmatic meanings
(5) Incongruent contexts lowered ratings for pragmatic meanings

And for the reaction-time data:

(6) Salient meanings were rated fastest, then nonsense, then pragmatic meanings
(7) Congruent contexts sped up reaction times for salient meanings; incongruent contexts slowed them down
(8) Congruent and incongruent contexts sped up reaction times for pragmatic meanings

In principle these are valuable points of comparison between typical adults and clinical or typically-developing samples. It is worth nothing though that the significant effects reported for ratings data typically had larger effect sizes than those for reaction-time data; variance was smaller for ratings data than for reaction-time data. Accordingly it would be unsurprising if with clinical and typically-developing populations the reaction-time effects were not replicated; we can reasonably expect greater variance in such samples. The following chapters will

outline an alternative way to compare group performance that was designed to probe beneath the considerable variance to find group differences.

Conclusions

This adult study satisfied its aims in the following ways: (1) it addressed the need for an experimental study that employs particularized conversational implicatures; (2) it adapted Surian et al's (1996) design in such a way that pragmatics was more clearly accessed; (3) it tested among models of utterance processing, finding most support for the graded-salience model and for the direct-access model; (4) it established various points of comparison for clinical and developmental groups.

The following chapters will move on to discuss adults with Autistic Disorder and two cohorts of typically-developing children. There is at this stage no mention of important limitations to the experimental design; given that the following studies follow the same experimental design as this adult study, discussion of such limitations is postponed to the general discussion chapter which follows the two empirical chapters.

Chapter 5: Study with Adults with AD/ASD

Introduction

This chapter describes a study on implicature in adults with Autistic Disorder (AD). Although AD is typically taken to cause pragmatic dysfunction, what this means in practice is unclear: is pragmatics dysfunctional as a whole or, if not, which types of pragmatic phenomena are impaired and why? The literature review pointed to equivocality between semantic and pragmatic impairment in some cases of apparent lower-level pragmatic dysfunction (Level (1)); the review also signalled fragmentary and inconsistent evidence for problems in level (2) and level (3) pragmatics; and, finally, it signposted apparent evidence that implicature, the centrepiece of higher-level pragmatics, was unaffected in an AD sample insofar as scalar implicatures were concerned (Pijnacker et al, 2009). This picture was supported by the naturalistic study, which found possible evidence for impairment to pragmatics at levels (1), (2) and (3), but an absence of evidence for impairment at level (4). The study reported in this chapter focused on implicatures that are unequivocally particularised, that is, that are less constrained than scalar implicatures and more clearly context-dependent. The aim was to search for dysfunction in higher-level pragmatics. Since no studies have yet demonstrated impairment relative to normal speakers in unambiguously inferential pragmatics (i.e. secondary pragmatics), studies in this area look especially urgent.

Method

Participants

16 adults with an autistic spectrum disorder gave informed consent and took part in the experiment. Participants' ages ranged from 16 years to 66 years; 12 were male and 4 were female. These adults were recruited through an established pool of participants, an autism support group, and the website of the National Autistic Society. All participants had a diagnosis of autism or Asperger's Syndrome; the 5 participants from the established pool had previously had their diagnoses confirmed by a speech-language therapist. In one case, a participant also had a diagnosis of semantic-pragmatic language disorder.

Materials

The materials were identical to those in the study on typical adults.

Procedure

The researcher asked participants to choose a place where they would feel comfortable taking part in an experiment. In practice this meant that, in most cases, the researcher visited participants in their home; in two cases the researcher visited participants in a quiet, neutral place; in two other cases the researcher hosted participants in the Department of Experimental Psychology.

The researcher engaged participants in unstructured conversation to encourage them to be responsive; when reasonably natural conversation was achieved, the researcher gave the same instructions as given to typical adults, allowing participants to ask questions. The researcher gave practice examples to ensure that participants understood the task. The researcher then read the materials aloud, and timed participants' responses on a stopwatch, giving breaks when necessary. A stopwatch was considered adequate to capture expected differences in response times while minimising intrusiveness and attendant discomfort for participants.

Results

The following table and figures represent descriptive statistics for the data for adults with ASD.

Table 5.1: Descriptive Statistics for Adults with ASD

Condition	ASD Ratings (n = 16)		RT (s)	
	M	SD	M	SD
Nonsense & Null	1.61	.57	4.03	2.88
Nonsense & Context	1.58	.48	3.64	2.56
Salient & Null	3.53	.44	3.25	2.29
Salient & Congruent Context	3.39	.53	3.77	2.72
Salient & Incongruent Context	3.2	.47	4.45	3.41
Pragmatic & Null	2.69	.68	4.24	2.86
Pragmatic & Congruent Context	2.78	.7	4.80	4.09
Pragmatic & Incongruent Context	2.56	.65	5.47	3.69

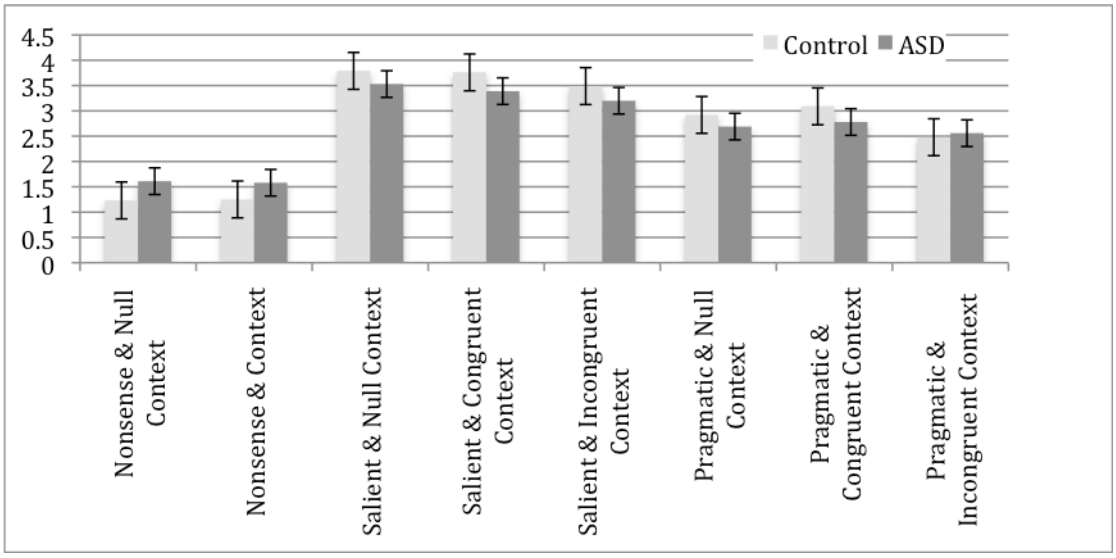


Figure 5.1: Typical adult controls versus adults with ASD - Average ratings by condition; error bars represent Standard Error

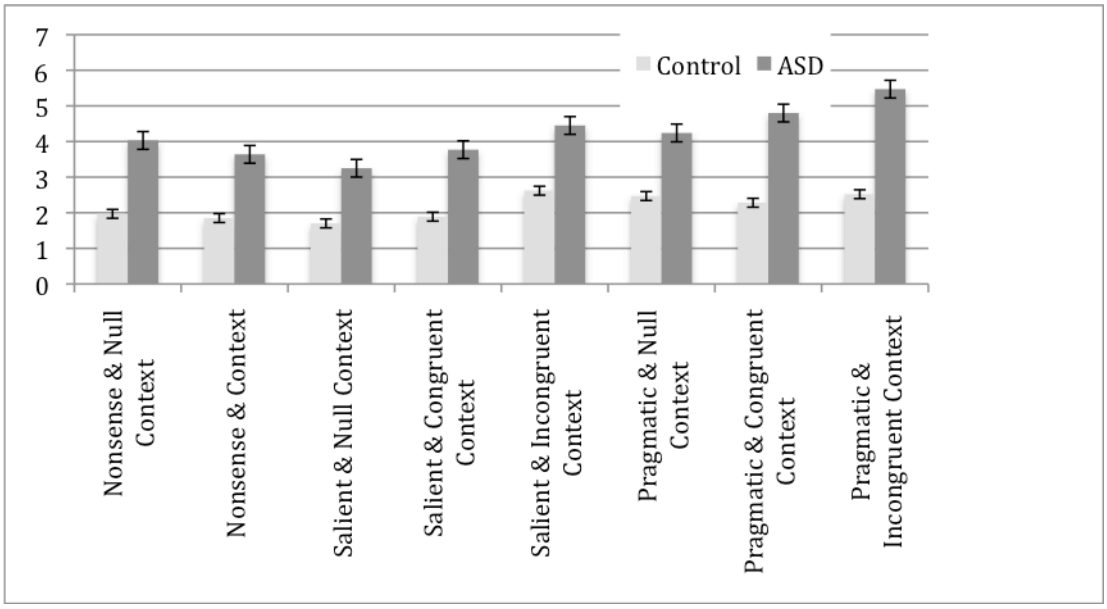


Figure 5.2: Typical adult controls versus adults with ASD - Average reaction time (s) by condition; error bars represent Standard Error

Ratings Analysis

Mean ratings were created for each participant by condition. Then for a simple and comprehensible analysis, two summary statistics were created on which statistical analysis was performed.

Composite Difference Scores

The first summary statistic corresponded to the variable Meaning Type. Typical adult data revealed that ratings for nonsense approximated to 1; ratings for pragmatic meanings approximated to 3; and ratings for salient meanings approximated to 4. These ideal scores were used as a benchmark for all individual participants' performance. A summary difference score was then produced for each participant, first, by subtracting their average rating for the nonsense condition from 1, for the pragmatic condition from 3, and for the salient condition from 4; second, by converting the products into absolute values; and third, by adding together the absolute values to produce a composite difference score. So, for example, a participant who rated nonsense as 1, pragmatic meanings as 3.5, and salient meanings as 3.7 would have absolute difference scores for each condition as 0, 0.5, and 0.3, and so a composite difference score of 0.8. Composite difference scores were calculated for all participants in the typical adult control group and in the ASD experimental group.

The composite difference scores for the control and experimental groups were then entered into an independent t-test with difference score as test variable and group (control, ASD) as grouping variable. Since Levene's test for equality of variances produced a non-significant statistic ($p = .43$), equal variances were assumed. The groups differed significantly ($t\ (1, 40) = 4.22, p < .05$)), with control participants producing lower composite difference scores than participants with ASD. The following table summarises the descriptive statistics for composite difference scores.

Table 5.2: Descriptive statistics for composite scores by group

	M	*SD*
ASD Group	3.46	1.27
Control Group	1.87	1.14

Context-sensitivity Scores

The second summary statistic corresponded to the variable Context Type. To assess how sensitive participants were to context, a context-sensitivity score was created for each participant by subtracting the mean rating in incongruent context from the mean rating in congruent context separately for each of salient and pragmatic meaning types.

The context-sensitivity scores for the typical-adult controls and the adults with ASD were entered into a one-way ANOVA, with Context Sensitivity with Salient Meaning and Context Sensitivity with Pragmatic Meaning as dependent variables and Group as a factor. There were no significant differences, or interactions, in context sensitivity. The following table summarises the descriptive statistics for context-sensitivity scores.

Table 5.3: Descriptive statistics for context-sensitivity scores by group

	Salient meaning		Pragmatic meaning	
ASD Group	.39	.34	.54	.32
Typical Adult Group	.39	.33	.4	.18

Analysis of the summary statistics identified a significant difference between the ASD and control groups in composite difference scores, a measure that corresponds

to the variable of Meaning Type. The analysis did not localise the difference more precisely; to do so, the ratings data were entered into mixed ANOVAS, firstly for the conditions without contexts, and secondly for those with contexts.

ANOVA for conditions without contexts

A mixed ANOVA was performed with Meaning Type (Nonsense, Salient, Pragmatic) as within-subjects variable and Group (Control, ASD) as between-subjects variable. The assumptions of sphericity and equal variance were satisfied. There was a significant effect of Meaning Type (F (2, 80) = 275.52, $p < .05$, partial $\eta 2 = .87$): pairwise comparisons revealed that nonsense was rated lowest (M = 1.43), pragmatic meanings next (M = 2.8), and salient meanings highest (M = 3.66); all differences were significant ($p < .05$). There was also a significant effect of Group (F (1, 40) = 181.23, $p < .05$, partial $\eta 2 = .82$) with the ASD group (M = 2.61) giving ratings significantly lower than the Control Group (M = 2.64). Finally, the interaction of Meaning Type and Group was significant (F (2, 80) = 7.01, $p < .05$, partial $\eta 2 = .15$). The following figure represents the interaction.

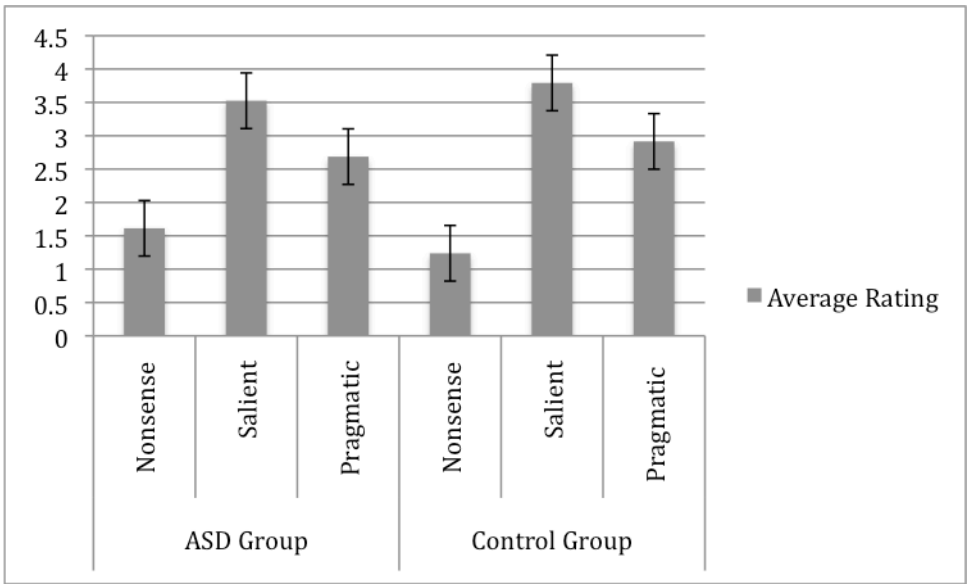


Figure 5.3: Typical adult controls versus adults with ASD - Ratings Data - Interaction of Meaning Type (Nonsense, Salient, Pragmatic) and Group (ASD, Control)

The Control Group showed a greater difference between nonsense and salient conditions than did the ASD group.

ANOVA for conditions with contexts

Another mixed ANOVA was performed, on the conditions with contexts, with Meaning Type (Salient, Pragmatic) and Context Type (Null, Congruent, Incongruent) as within-subjects variables and Group (ASD, Control) as between-subjects variable. The assumption of sphericity and equal variance were satisfied. There was a significant main effect of Meaning Type (F (1, 40) = 149.87, $p < .05$, partial $\eta^2 = .79$): salient meanings ($M = 3.53$) were rated significantly higher than pragmatic meanings ($M = 2.75$). The main effect of Context Type was also significant (F (2, 80) = 14.29, $p < .05$, partial $\eta^2 = .26$): the differences between null and incongruent contexts, on the one hand, and congruent and incongruent contexts, on the other were significant; the following table summarises the descriptive statistics.

Table 5.4: Descriptive statistics for Context Type variable

Condition	*M*
Null	3.23
Congruent	3.25
Incongruent	2.93

The main effect of Group was also significant (F (1, 40) = 4.08, p < .05, partial η^2 = .09) with the ASD group giving lower ratings (M = 3.02) than the Control Group (M = 3.26). None of the interactions was significant, suggesting that the differences in the conditions with contexts were global rather than localised.

Reaction-time Analysis

Because they were not normally distributed, the reaction-time data were normalised using the natural log of each value.

Without Contexts

A mixed ANOVA was performed, with Meaning Type (Nonsense, Salient, Pragmatic) as within-subjects variable and Group (Control, ASD) as between-subjects variable. The assumptions of sphericity and equal variance were satisfied. There was a significant main effect of meaning type (F (2, 39) = 9.14, p <.05, partial η2 =.19), with pairwise comparisons revealing that reaction times were significantly faster (p < .05) for nonsense than pragmatic meanings, and for salient meanings than pragmatic meanings. There was also a significant main effect of Group (F (1, 40) = 11.67, p <.05, partial η^2 =.23), with controls' average reaction time (M = .55) being significantly faster than adults' with ASD (M = 1.07). The interaction of Meaning and Group was marginally significant (F (2, 80) = 2.94, p = .06, partial η^2 = .07). The following figure represents the interaction.

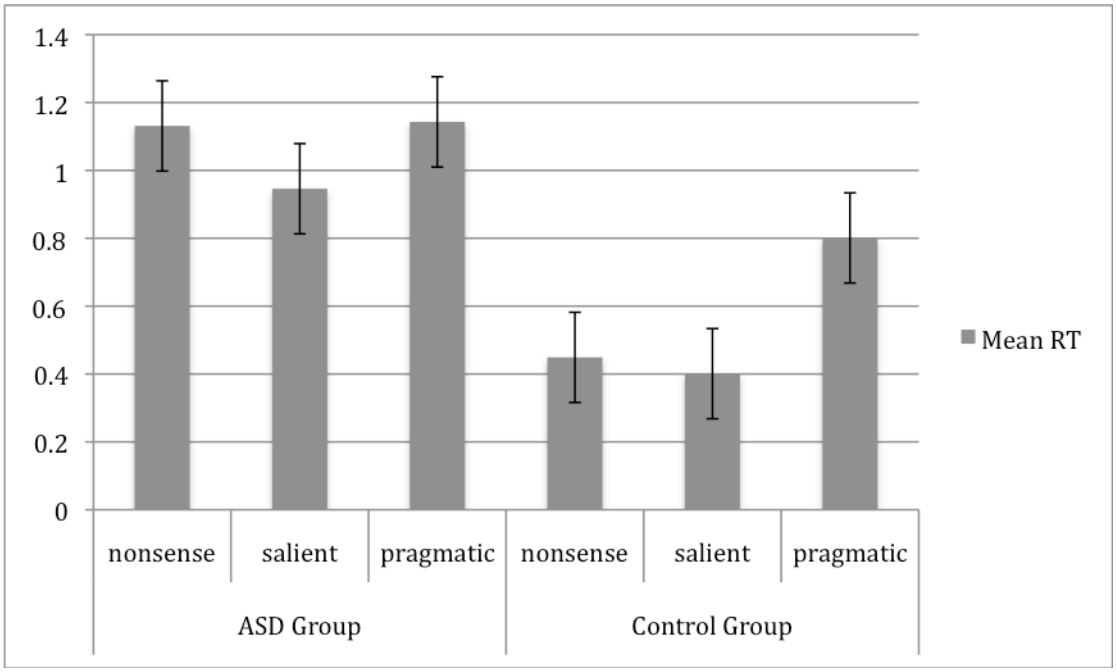


Figure 5.4: Typical adult controls versus adults with ASD - Reaction-time Data - Interaction of Meaning Type (Nonsense, Salient, Pragmatic) and Group (ASD, Control)

Whereas, in the control group, pragmatic meanings take considerably longer to process than nonsense (taking .4s for controls to rate nonsense and .8s to rate pragmatic meanings), in the ASD group the difference is slight (taking 1.13s to rate nonsense and 1.14s to rate pragmatic meanings).

Reaction Times With Context

Another mixed ANOVA was performed, this time with Meaning Type (salient, pragmatic) and Context Type (null, congruent, incongruent) as within-subjects variables and Group (control, ASD) as between-subjects variable. Again, the assumptions of sphericity and equal variance were satisfied. There was a significant main effect of Meaning Type (F (1, 40) = 23.72, p <. 05, partial η^2 =. 37): reaction times for salient meanings (M = .80) were faster than those for pragmatic meanings (M = 1.02). There was also a significant main effect of Context Type (F (2, 80) =

12.40, $p < .05$, partial $\eta^2 = .24$): meanings in null contexts were processed fastest ($M = .82$), then meanings in congruent contexts ($M = .88$), with the difference in processing between the two being non-significant; meanings in incongruent contexts were processed significantly slower than both other conditions ($M = 1.03$, $p < .05$). There was also a significant main effect of Group (F (1, 40) = 9.52, $p <$ .05, partial $\eta^2 = .19$): control reaction times ($M = .66$) were significantly faster than ASD reaction times ($M = .66$). The interaction of Meaning Type, Context Type, and Group was also significant (F (2,80) = 4.66, $p <$. 05, partial $\eta^2 = .1$).

Discussion

The data outlined above offer interesting evidence of impairment to higher-level (level (4)) pragmatics in adults with AD/ASD. Composite difference scores were significantly higher for the ASD experimental group than for the control group; hence adults with ASD deviated significantly more from the ideal ratings of 1 for nonsense, 3 for pragmatic meanings, and 4 for salient meanings. The ASD group did, however, favour pragmatic meanings to nonsense; since these pragmatic meanings were meaningful in virtue of an underlying implicature, the ASD group showed at least a basic sensitivity to implicature in preferring the pragmatic meanings to nonsense. Intriguingly, there was no significant difference between the ASD and control groups in sensitivity to context as measured by the context-sensitivity summary statistic. The ASD group was significantly slower in delivering ratings than was the control group.

The lack of a significant difference in context sensitivity may be misleading. Certainly it is striking that the ASD group is not simply blind to context, and that participants with ASD accordingly made some use of the contexts provided. But it is one thing to make use of context; it is quite another to make use of an ostensive context in the appropriate way. Witness the performance of Participant 16, who seemed hesitant to use the contexts provided but offered a highly idiosyncratic context of his own for the following nonsensical item:

Q: What would you like for dinner?
A: Green grass covered in mud

For this item, Participant 16 suggested that if the speaker in A were a child who liked using pet names for certain foods, he could have meant cabbage (green grass) covered in gravy (mud). Because the structure of the experimental task did not encourage participants to 'show their working', this explicit rationalisation was an isolated case. Future studies could encourage participants to rationalise their choices, though at the risk of creating heavy meta-linguistic demands. In any case, the intriguing possibility arises that people with ASD are not taxed by contexts *per se* but by using an ostensive context in an appropriate way to reach a particular implicit intended meaning.

The question naturally arises: What determines the appropriate way to use a context to reach a particular implicit intended meaning? Many answers can be found in pragmatic theory. A hearer might, for instance, have to generate hypothetical meanings in order of accessibility, stopping when an expectance of relevance is met, that is, when the processing demands are offset by the cognitive effects gained (e.g. derivable true conclusions). To probe such accounts further, we would need to

apply pragmatic theories to experimental research, perhaps, for example, extending the work of Van der Henst and Sperber (2004) in testing principles of Relevance Theory with relational-judgement tasks, the Wason Selection task or time-telling tasks.

The only existing study to probe implicature in adults with ASD was conducted by Pijnacker et al (2009), who found no significant difference between their ASD participants and their control participants overall, although their participants with high-functioning autism did differ significantly from their participants with Asperger's. In contrast, the present study has found significant differences between typical adults and adults with ASD on an implicature task. These results suggest impairment in level (4) pragmatics. The suggestion of such impairment sits well with existing data suggesting impairment to metaphor and irony (Happé, 1993).

The difference between the present findings and those of Pijnacker et al (2009) deserves explanation. Differences between the studies can be explored using a framework from Pouscoulous et al (2007). In a study with various age groups, to be described more extensively in the next chapter, Pouscoulous et al explored children's performance on a comprehension task with scalar implicatures by varying the task demands. In particular the authors varied the type of task, the type of scale used (positive or negative) and the number of filler items, finding that meta-linguistic tasks, negative scales (and perhaps by extension negative implicatures) and the presence of filler items all reduced the number of implicatures produced. Both the present study and Pijnacker et al used a meta-linguistic task (ratings of acceptability and truth-value judgement respectively), and so should in this respect

have tasks of more-or-less equivalent difficulty. Both studies also used affirmative implicatures, and so should in this respect be equivalent. While Pijnacker et al used filler items, the present study did not and so in this respect should have an easier task. The most likely causes of the difference, then, are either a difference in level of functioning or the different type of implicatures used. The former possibility cannot be pursued because the present study did not include measures of intelligence; the latter possibility has been discussed in preceding chapters and so will not be discussed at great length here. Suffice it to say, to explore differences between types of implicature, exactly equivalent methods could be used.

The final point to be made in this chapter is that although a significant difference has been found between the ASD and control groups, this difference could either signal a developmental delay or an eccentric path of development. To try to answer the question of which option is most likely, this thesis will compare adults with ASD with two groups of typically developing children.

Chapter 6: Typically Developing Children

Introduction

This child study was doubly motivated. The foregoing chapter on the study with adults with ASD gave the first motivation: the study reported in that chapter found significant differences between the Control and ASD group; but that difference was hard to interpret without developmental data to test whether the ASD performance showed a developmental delay or a thoroughly eccentric trajectory. The second motivation is that the existing developmental literature focuses on a subset of implicatures, the scalar implicature, leaving much territory unexplored. The studies with typical adults and with adults with ASD both focused on less constrained implicatures, and thus offer a methodology for exploring implicatures more generally with children.

Preceding chapters have already outlined the details of the experiment and the motivation to test various models of utterance processing; this brief introduction will sketch the developmental literature on the implicature, which, like the adult literature, focuses on scalar implicature. The literature on scalar implicatures is young. There were hints in research on the acquisition of quantifiers (Smith, 1980) and propositional connectives (Braine & Rumain, 1981) that young children are not competent with implicature. But the first experimental work dedicated to implicature was conducted only recently (Noveck, 2001). Noveck (2001) presented children aged 5, 7 and 9 and adults with reasoning tasks manipulating the modal verbs ‘must’ and ‘might’ and the French quantifiers ‘certains’ (some) and ‘tous’ (all). Noveck’s participants performed truth-value judgements on a full array of

control sentences and on experimental sentences that were logically false (e.g. ‘There must be a parrot in the box’ when there need not be), logically true but pragmatically infelicitous (e.g. ‘There might be a parrot in the box’ when there had to be), and logically true and pragmatically felicitous (e.g. ‘There must be a parrot in the box’ when there indeed had to be). Any pragmatically competent participant should reject logically true but pragmatically infelicitous sentences. With the modal verbs, the youngest of Noveck’s participants to perform the task well were 7-year-olds. Of the 7-year-olds, 80 % accepted logically true but pragmatically infelicitous sentences; of the 9-year-olds, 69% accepted these sentences; but of the adults, only 35% accepted them. With the quantifiers, only adults rejected logically true but pragmatically infelicitous sentences. Noveck took his data to show that children behave more logically but less pragmatically than children.

Noveck’s findings have been extended and refined by various researchers. Papafragou and Musolino (2003) replicated Noveck’s finding that children were less likely to recover an implicatum, and extended the finding to other scales. They used the English quantifier ‘some’ (which implies ‘not all’), the numeral ‘three’ (which, like other numerals, implies ‘exactly that amount’), and the encyclopaedic scale ‘start’. In an encyclopaedic scale, an implicature arises from a stable relationship between concepts. So for ‘start’, if a student says, ‘I have started writing my essay’ when he has in fact finished, his utterance is pragmatically infelicitous, because he could more informatively have said that he had finished. Papafragou and Musolino (2003) abandoned truth-value judgements, told participants that they would hear a puppet saying something ‘silly’ and that they should help the puppet say it better, and included contexts that alerted participants

to the clash between logical and pragmatic readings. They also trained participants. In so doing the authors elicited better performance in 5-year-olds than Noveck did, pushing the age of acquisition somewhat lower. Papafragou and Tantalou (2004) extended such findings still further by testing a broader range of implicatures, and by providing more vivid, natural contexts, in which participants had to withhold prizes from characters who provided under-informative statements, and rationalise their choice. They also trained their participants. Their participants, aged between 4 years and 1 month and 6 years, consistently performed well on the task and gave coherent rationales. These findings held for all types of scalar implicature; there were no significant differences among the types. The only difference was that rationales were more variable for *ad hoc* scalars than for the other types. Overall, then, there are robust findings that children are more logical, and less pragmatic, than adults in interpreting scalar terms; but the age at which children perform pragmatically in a test can be remarkably young depending on the method used.

The foregoing studies draw a complex picture of the implicature; in a more recent study, which was discussed briefly in the preceding chapter, Pouscoulous, Noveck, Politzer and Bastide (2007) explored various factors that might explain differences in participants' behaviour. The authors conducted a series of experiments that manipulated the number of distracters in the materials, the type of task, the choice of scalar expression, and the type of scale. They found that the following factors improved performance: having fewer distracters; having as a dependent measure an acted-out response rather than a meta-linguistic judgement; and using an affirmative, rather than a negative, scale. The choice of scalar expression also influenced performance: children were more likely to produce an implicature with

the quantifier ‘quelques’ than with the quantifier ‘certains’; adult production did not differ between the quantifiers. All supporting factors encouraged pragmatic performance in children, so that children as young as 4 could successfully produce implicatures.

Because Pouscoulous et al (2007) published their study after data collection had begun for this thesis, their findings could not be fully incorporated into the present design. However, we will find it useful to consider, here, how their findings relate to the present design. Pouscoulous et al suggest minimising the number of distracters to encourage children to respond pragmatically. In the present study all items were either critical pragmatic items (i.e. those that depended on an implicature) or essential controls; there were no additional distracters. While the number of conditions could have been reduced, for example by presenting items only with congruent contexts, the reduction would have created asymmetry with the typical adults, and would have reduced the number of points of comparison. Pouscoulous et al suggest that having participants act out their responses, rather than make them in meta-linguistic judgements, encourages them to respond pragmatically because meta-linguistic judgements are more taxing: meta-linguistic tasks presuppose not just linguistic abilities but also insight into these abilities. Acting-out strategies can readily be incorporated into experiments on expressions, such as scalar implicatures, that engender a constrained set of responses. But if the research focus is on more open-ended inferences, as with other types of implicature, then acting-out strategies become much harder to apply; accordingly there may yet be a place for meta-linguistic judgements, and especially if, like the judgements in the present task, they allow subtle comparisons between conditions. Pouscoulous et

al, finally, suggest that affirmative scales are easier to process; the present study overwhelmingly used affirmative implicatures.

Existing developmental studies on the implicature have a focus that is more descriptive than explanatory: that is, their motivation is to define the developmental trajectory for implicatures; they give relatively little attention to what explains the trajectory. Although this focus is justifiable (we can hardly explain a phenomenon if we do not know what it is), research could profit from testing possible explanatory mechanisms; this study attempts to add such a test. This thesis has already rejected the use of theory of mind as a complete explanation of pragmatic abilities. An alternative would be to test more local, linguistic explanations, that is, to test the competing pragmatic theories. One such theory, Relevance Theory, is increasingly tested experimentally: Van der Henst and Sperber (2004), for instance, have tested relevance-theoretic claims about cognition and communication using relational-judgement tasks, the Wason Selection Task, and time-rounding tasks with digital and analogue clocks. As yet, though, these methods have not been generalised to children. This study instead tested the models of utterance processing that were tested in the preceding adult study.

A related question is whether models of utterance processing imply a cognitive structure that is separate from central processing, or whether they instead imply that pragmatics depends on cognition more generally. The question pertains to the philosophical debate into the modularity of mind, in which pragmatics has begun to feature quite strongly (e.g. Sperber & Wilson, 2002). The question also pertains to discussion of pragmatic dysfunction. A straightforward line of reasoning might run

as follows: if pragmatics is found to depend on intelligence in typically developing speakers, then pragmatic dysfunction may be explicable in broader cognitive terms; if pragmatics is modular, however, more localized explanations must be found. There are grounds to suppose that pragmatic skills are more-or-less selectively impaired in autistic spectrum disorder, the primary case of pragmatic dysfunction: very able speaker-hearers with ASD show high intelligence but low pragmatic skills. Investigating an association between pragmatics and intelligence might seem unprofitable in light of such dissociation. We might nevertheless still find an association between pragmatics and intelligence in the typical population; that is, typical speakers and speakers with ASD might show different patterns of encapsulation, with typical speakers being able to deploy general intelligence to help in pragmatics tasks and speakers with ASD being unable to do so. Since pragmatic skills arguably contribute to tests of verbal intelligence[18], a more general test was selected, namely Ravens Coloured Progressive Matrices. Raven's Matrices gave a quick and reliable estimate of participants' general intelligence. A standardized test of language was not used, for the following reasons. While the ability to recover implicatures does depend on a certain level of language function, it requires more sophisticated inferential cognition that a grammatical test would not tap into. The nonsense condition is as much of a language test as is required for current purposes.

[18] The Similarities Test (e.g. 'How are a snake and an alligator alike?') is arguably, for instance, a test of *ad hoc* concept construction, in which concepts are pragmatically calibrated for immediate conversational purposes.

Method

Design

The same design was used here as in the previous experimental chapters.

Participants

41 typically developing children took part in the study. These children formed two cohorts. Children in the first cohort were in Year 3 classes; they were aged between 7 and 8. Children in the second cohort were in Year 5 classes; they were aged between 9 and 10. These age-groups were chosen because pragmatic performance has reliably been shown to begin to emerge by about 7-years' of age; the younger groups, used for example by Papafragou and Musolino (2003) and Papafragou and Tantalou (2004), were avoided because pilot testing showed that younger children struggled with the more complex task format required by the type of implicature used. The first cohort was included to probe the earliest stages of implicature development; the second, to track the development of implicature. Children were recruited through state primary schools in rural Oxfordshire. Informed consent was sought from parents using an opt-in procedure. In the first cohort, 10 boys and 10 girls were recruited. The average age was 7 years and 5 months (range 6; 11 – 8;0); the average score on Ravens matrices was 107.35. In the second cohort, 11 girls and 10 boys were recruited. The average was 9 years and 5 months (range 8; 0 -10; 7); the average score on Ravens matrices was 110.95. The groups did not differ in average score on Ravens Matrices using age-scaled scores ($t(1,19) = 1.03$, $p = .32$, *2*-tailed). All children were native speakers of English; none had a history of neurological disorders, or developmental or acquired language disorders.

Materials

Pilot testing suggested certain modifications to the materials. When the materials were simply real aloud by the experimenter, as they had been to normal adults and adults with autism, the children struggled to maintain their attention. To keep the children's attention, the micro-conversations were acted out on video with a frog puppet and a rabbit puppet, named Mr Frog and Mr Rabbit respectively. Mr Frog always asked the question; Mr Rabbit always responded. Items were also reworded, wherever necessary, to avoid difficult vocabulary.

Procedure

Experimental sessions took place in an empty classroom in the participant's school. To relax the participant and encourage him/her to be more talkative, there was first some unstructured conversation. The experimental session proper then began. The session comprised two tasks, namely the implicature experiment and Raven's Coloured Progressive Matrices (i.e. Sets A, AB and B). The order in which these tasks were presented was counterbalanced. For the Ravens Coloured Progressive Matrices, the standardized format was followed. For the implicature experiment, the following instructions were given. Children were told that Mr Frog would ask Mr Rabbit lots of questions; that Mr Rabbit would try to answer Mr Frog's questions; and that sometimes Mr Rabbit would make sense and sometimes he wouldn't. They were then told that, if Mr Rabbit said something nonsensical, they should give him 1; if he said something not very good, they should give him 2; if he said something ok, they should give him 3; and if he said something that made perfect sense, they should give him 4. They were told that they would be timed but that they shouldn't race. The length of time between the last word of the micro-conversation and the

first word of the participant's response was recorded. Participants were given a few practice questions so that they felt comfortable with the drill and so that they demonstrated an understanding of the instructions. Between tasks children were given a short break in which there was, again, some unstructured conversation. To maximize participants' performance, the experimental session was kept as short as possible and, in any case, shorter than 30 minutes.

Analysis

The following tables summarize the descriptive statistics.

Table 6.1: Descriptive statistics for ratings data

Condition	Adults with ASD (N = 16)		Typical Adults (N = 26)		Younger Children (N = 20)		Older Children (N = 21)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Nonsense & Null	1.61	.57	1.24	.4	1.45	.49	1.28	.30
Nonsense & Context	1.58	.48	1.25	.46	1.48	.41	1.22	.25
Salient & Null	3.53	.44	3.79	.24	3.61	.42	3.37	.52
Salient & Congruent	3.39	.53	3.76	.33	3.56	.36	3.42	.51
Salient &	3.2	.47	3.49	.42	3.48	.38	3.14	.58

In-congruent								
Pragmatic & Null	2.69	.68	2.92	.56	2.77	.60	2.50	.53
Pragmatic & Congruent	2.78	.7	3.09	.58	2.86	.71	2.56	.77
Pragmatic & In-congruent	2.56	.65	2.48	.67	2.7	.56	2.27	.68

Table 6.2: Descriptive statistics for reaction-time data

Condition	Adults with ASD (N = 16)		Typical Adults (N = 26)		Younger Children (N = 20)		Older Children (N = 21)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Nonsense & Null	4.03	2.88	1.97	1.92	3.95	2.17	2.29	1.57
Nonsense & Context	3.64	2.56	1.85	1.73	3.72	3.46	1.77	.87
Salient & Null	3.25	2.29	1.7	1.02	9.36	3.95	1.95	1.33
Salient & Congruent	3.77	2.72	1.89	1.39	3.66	4.26	1.9	.95
Salient &	4.45	3.41	2.62	1.8	5.22	3.42	1.96	.83

In-congruent								
Pragmatic & Null	4.24	2.86	2.47	1.23	4.56	3.42	2.13	1.5
Pragmatic & Congruent	4.80	4.09	2.28	1.11	4.28	3.4	2.52	2.33
Pragmatic & In-congruent	5.47	3.69	2.52	1.32	3.83	4.89	2.08	1.22

Ratings

As in the preceding chapter, mean ratings were created for each participant by condition, and the two summary statistics were calculated and entered into further analyses. Here, for the sake of completeness, analyses including all groups are reported.

Composite Difference Scores

Composite difference scores were created for all groups as described in the preceding chapter. Recall that these scores correspond to the Meaning Type variable. The scores for all groups were entered into a one-way ANOVA with score as dependent variable and Group as factor. Levene's test for homogeneity of variances produced a non-significant statistic, so homogeneity of variances was

assumed. The effect of Group was significant (F (3, 82) = 7.41, p <.05). The following table summarises the descriptive statistics for the data.

Table 6.3: Descriptive statistics for composite difference scores by group

	M	*SD*
Young Children	2.91	.96
Older Children	2.99	.1.28
ASD Group	3.46	1.27
Typical Adult Group	1.87	1.14

Post-hoc tests with the Bonferroni Adjustment for multiple comparisons showed the following results. The following differences were significant: between the Typical Adult Group and the ASD group (p < .05); between the Typical Adult Group and the Young Children Group (p < .05); and between the Typical Adult Group and the Older Children Group (p < .05). The following differences were non-significant: between the ASD Group and the Young Children (p = .97); between the ASD Group and the Older Children (p = 1.0); between the Young Children and the Older Children (p = 1.0).

Context-sensitivity scores

Context-sensitivity scores were calculated as described in the preceding chapter for all groups of participants. The scores were then entered into one-way ANOVAs, one with context-sensitivity score with salient meanings as dependent variable and one with context-sensitivity score with pragmatic meanings as dependent variable;

in both cases, Group was the factor. There were no significant differences between the groups with either salient or pragmatic meanings.

Table 6.4: Descriptive statistics for context-sensitivity scores by group

	Salient meaning		Pragmatic meaning	
	M	*SD*	*M*	*SD*
Young Children	.4	.18	.6	.46
Older Children	.56	.41	.57	.41
ASD Group	.39	.34	.54	.32
Typical Adult Group	.39	.33	.4	.18

To try to localise the differences between groups that were detected in the composite difference scores, the ratings data were entered into the following mixed ANOVAs.

ANOVA for conditions without contexts

A mixed ANOVA was run on the ratings data with Meaning Type (Nonsense, Salient, Pragmatic) as within-subjects variable and Group (ASD, Typical Adult, Young Children, Older Children) as between between-subjects variable. The assumptions of sphericity and equal variance being satisfied, the ANOVA produced the following results. The main effect of Meaning Type was significant (F (2, 158) = 527.89, p <. 05, partial η^2 = .87): pairwise comparisons revealed that differences between all conditions were significant (p < .05), with nonsense rated lowest (M = 1.4), pragmatic meanings next (M = 2.7) and salient meanings highest (M = 3.6). The main effect of Group was also significant (F (3, 79) = 2.87, p < .05, partial η^2 = .09): pairwise comparisons suggested that only the difference between the Typical

Adult Group and the Older Children Group was significant ($p < .05$). The interaction of Meaning Type and Group was significant (F (6, 158) = 2.78, $p < .05$, partial $\eta^2 = .1$).

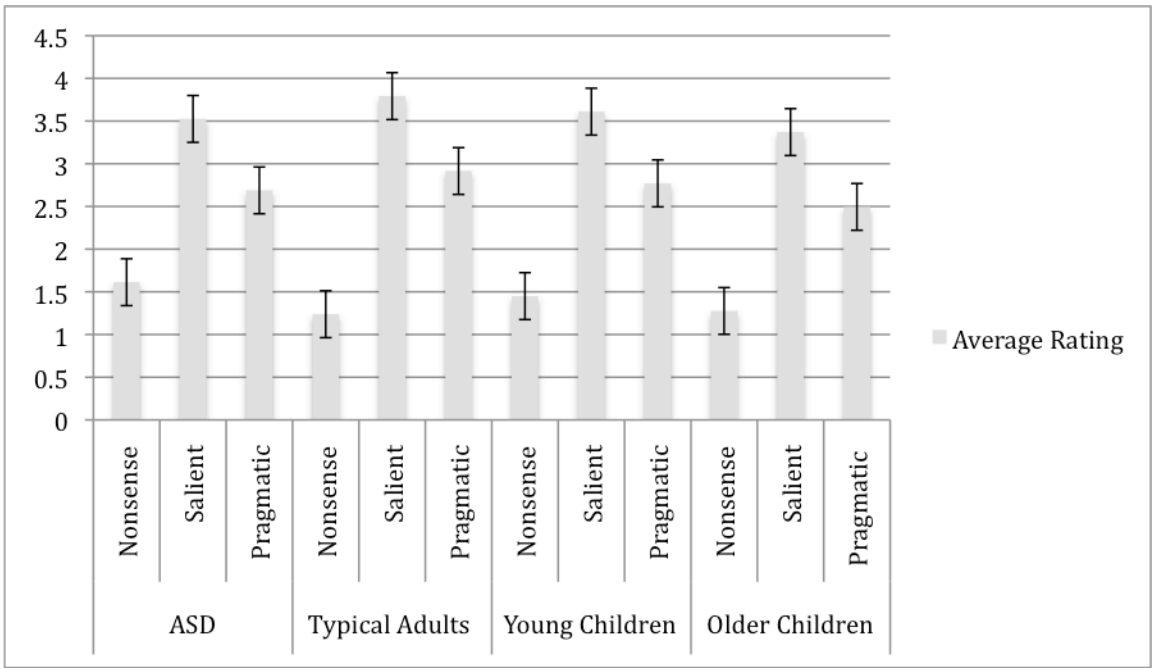


Figure 6.1: All participants - Ratings Data - Interaction of Meaning Type and Group

ANOVA for conditions with contexts

Another mixed ANOVA was performed, this time with Meaning Type (Salient, Pragmatic) and Context Type (Null, Congruent, Incongruent) as within-subjects variables and Group (ASD, Typical Adult, Younger Children, Older Children) as between-subjects variable. The main effect of Meaning Type was significant (F (1, 79) = 348.37, $p < .05$, partial $\eta^2 = .82$): salient meanings ($M = 3.48$) were significantly preferred ($p < .05$) to pragmatic meanings ($M = 2.68$). The main effect of Context Type was also significant (F (2, 158) = 15.34, $p < .05$, partial $\eta^2 = .16$): congruent contexts were rated highest ($M = 3.18$), closely followed by null contexts ($M = 3.15$), followed by incongruent contexts ($M = 2.92$); the differences between congruent contexts and incongruent contexts and between null contexts and

incongruent contexts were significant ($p < .05$). The main effect of Group was also significant (F (1, 79) = 4.48, $p < .05$, partial $\eta^2 = .15$): only the difference between the Typical Adult Group and the Older Children Group was significant ($p < .05$). No interaction was significant.

Reaction Times

Because the reaction-time data were not normally distributed, they were normalised using the natural log of each value.

ANOVA for conditions without contexts

A mixed ANOVA was run with Meaning Type (Nonsense, Salient, Pragmatic) as within-subjects variable and Group (ASD, Typical Adults, Young Children, Older Children) as between-subjects variable. The assumptions of sphericity and of equal variance were satisfied. The main effect of Meaning Type was significant (F (1, 79) = 4.32, $p < .05$, partial $\eta^2 = .09$): salient meanings were rated fastest ($M = .7$), then nonsense ($M =. 8$), then pragmatic meanings ($M = .91$); only the difference between salient meanings and pragmatic meanings was significant ($p < .05$). The main effect of Group was also significant (F (1, 79) = 5.22, $p < .05$, partial $\eta^2 = .16$. Timings were as follows:

Table 6.5: Mean reaction time by group

Group	Mean Reaction Time	Standard Error
ASD Group	1.07	.12
Typical Adults	.55	.09

Young Children	.94	.11
Older Children	.66	.1

Pairwise comparisons revealed that the differences between the ASD Group and the Typical Adult Group, between the ASD Group and the Older Children Group, and between the Typical Adult Group and the Younger Children Group were all significant ($p < .05$).

The interaction of Meaning Type and Group was also significant (F (6, 158) = 2.34, $p < .05$, partial $\eta^2 = .08$); the following figure represents the interaction.

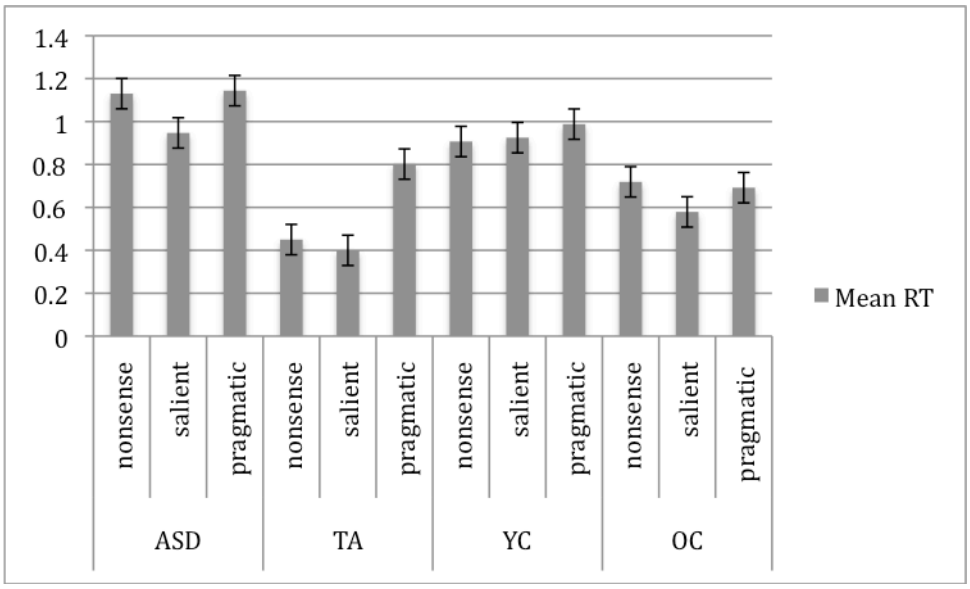


Figure 6.2: All participants - Reaction-time Data - Interaction of Meaning Type and Group

Whereas the ASD Group, the Typical Adult Group and the Older Children Group all show the pattern of salient > nonsense > pragmatic, the Younger Children Group show the pattern nonsense > salient > pragmatic.

ANOVA for conditions with contexts

A mixed ANOVA was run with Meaning Type (Salient, Pragmatic) and Context Type (Null, Congruent, Incongruent) as within-subjects variables and Group (ASD,

Typical Adults, Young Children, Older Children) as between-subjects variable. The assumptions of sphericity and equal variance were satisfied. The main effect of Meaning Type was significant (F (1, 79) = 35.05, $p < .05$, partial $\eta^2 = .31$): salient meanings ($M = .76$) were rated significantly ($p < .05$) faster than pragmatic meanings (.94). The main effect of Context Type was also significant (F (2, 158) = 6.05, $p < .05$, partial $\eta^2 = .07$): null contexts were fastest ($M = .81$), then congruent contexts ($M = .83$), then incongruent contexts ($M = .91$); the differences between null and incongruent contexts and between congruent and incongruent contexts, but not between null and congruent contexts, were significant. The main effect of Group was also significant (F (1, 79) = 4.26, $p < .05$, partial $\eta^2 = .14$). Timings were as follows

Table 6.6: Mean reaction time by group

Group	Mean Reaction Time	Standard Error
ASD Group	1.16	.13
Typical Adults	.66	.1
Young Children	.91	.11
Older Children	.66	.11

Only the differences between the ASD Group and the Typical Adult Group and between the ASD Group and the Older Children Group were significant. The interaction of Context Type and Group was significant (F (6. 158) = 3.55, $p < .05$, partial $\eta^2 = .12$. The following figure represents this interaction; it suggests that while the Typical Adult Group and the Younger Children Group processed context types in the order congruent>null>incongruent, the ASD Group and the Older Children processed contexts in the order null>congruent>incongruent.

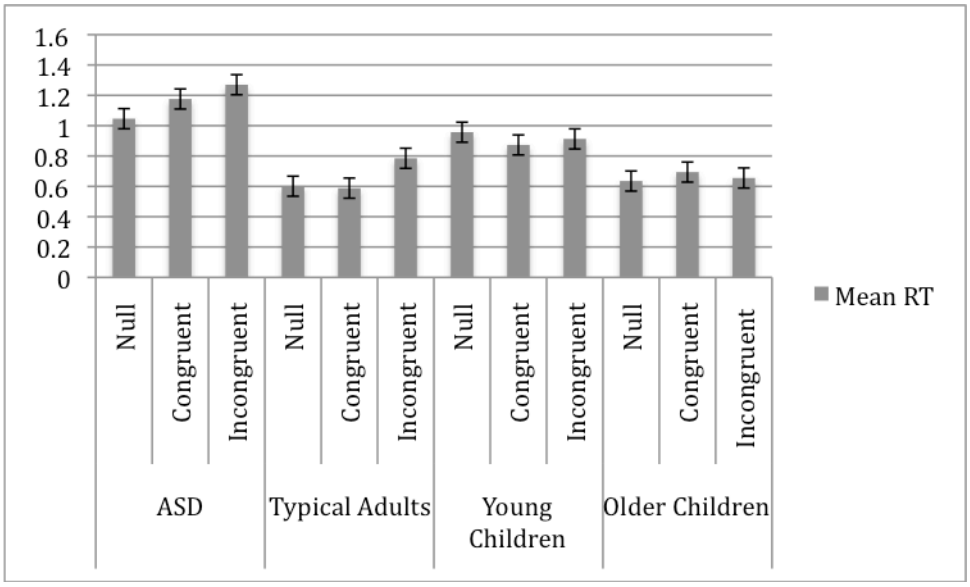


Figure 6.3: All participants - Reaction-time Data - Interaction of Context Type (Null, Congruent, Incongruent) and Group (ASD, Typical Adults, Young Children, Older Children)

The interaction of Meaning Type, Context Type and Group was also significant (F (6, 158) = 2.64, p < .05, partial η^2 = .09).

Models of utterance processing: child data only

Prediction Set A

A repeated-measures ANOVA was performed on data for items without context with Meaning Type (Nonsense, Salient, Pragmatic) as independent variable and rating as dependent variable. There was a significant main effect of Meaning Type: (F (2, 39) = 258.09, p <.05, partial η^2 = .87: pairwise comparisons revealed that participants significantly preferred pragmatic meanings to nonsense (p <.05) and salient meanings to pragmatic meanings (p <.05).

Predictions sets B and C: Meanings with Context

General Analysis

A 2 x 3 x 2 mixed ANOVA was performed on these data, with average rating as the dependent variable, Meaning Type (salient, pragmatic) and Context Type (null, congruent, incongruent) as within-subjects variables and Task Type (sense, foreign speaker) as a between-subjects variable. The assumption of sphericity was satisfied for all variables and their interactions, as was the assumption of equal variances. There was a significant main effect of Meaning Type ($F(1,40) = 206.14$, $p < .05$, partial $\eta^2 = .84$), with pairwise comparisons revealing that salient meanings were significantly preferred to pragmatic ones ($p < .05$). There was also a significant main effect of Context ($F(2, 39) = 3.75$, $p < .05$, partial $\eta^2 = .09$), with pairwise comparisons revealing that the only significant difference was between congruent and incongruent contexts. The interaction of Meaning Type and Context was non-significant ($F(2, 39) = .18$, $p = .84$).

To investigate the predictions of the models of utterance processing further, paired-sample t-tests were performed.

Prediction Set B

Of meanings in congruent contexts, salient meanings were rated significantly better than pragmatic ones ($t(40) = 6.99$, $p < .05$). Of meanings in incongruent contexts, salient meanings were rated significantly better than pragmatic ones ($t(40) = 9.40$, $p < .05$).

Prediction Set C:

With respect to salient meanings, congruent contexts made no significant difference to ratings ($t(40) = .14$, $p = .989$); incongruent contexts significantly lowered ratings ($t (40 = 2.28$, $p < .05$).

Prediction Set D:

With respect to pragmatic meanings, congruent contexts made no significant difference ($t (40) = .64$, $p = .52$) and nor did incongruent contexts ($t (40) = 1.23$, $p =$ 2.3.

Reaction-time predictions

Of the comparisons mandated by predictions from the models of utterance processing, only one was significant: a paired-sample t-test showed that salient meanings in congruent contexts were processed significantly faster than pragmatic meanings in congruent contexts ($t (1,40) = 3.89$, $p < .05$).

Pragmatic Competence and Intelligence

To test whether score on Ravens Progressive Matrices predicted pragmatic competence, data for participants' IQ and composite difference scores were analysed using a linear regression. IQ did not significantly predict composite difference score (ß = .09, $p = .39$), accounting for only .8% of the variance. Linear regressions were likewise run to test whether IQ predicted context-sensitivity scores for salient or pragmatic meanings. IQ did not significantly predict context-

sensitivity scores for salient meanings (ß = .15, p = .34), accounting for 2% of the variance; nor did IQ significantly predict context-sensitivity scores for pragmatic meanings (ß =-.02, p =.92), accounting for none of the variance.

A test of basic pragmatic competence

It was deemed advisable to test further whether the groups of participants all achieved a basic level of competence, by testing whether participants could accept inferences that in other contexts would be infelicitous, so across the groups mean ratings for pragmatic meanings with congruent contexts were compared with those for pragmatic meanings with incongruent contexts (Ann Dowker & Napoleon Katsos, 2010, *personal communication*). A mixed repeated-measures ANOVA made the relevant comparison; the within-subjects variable was Meaning Type (Pragmatic & Congruent, Pragmatic & Incongruent) and the between-subjects variable was Group (typical adults, ASD, young children, older children). The Bonferroni Adjustment for multiple comparisons was used, and the assumptions were satisfied. The main effect of Meaning Type was significant (F (1, 79) = 20.94, p <.05, partial η^2 =. 21): Pragmatic & Congruent meanings (M = 2.82) were significantly preferred to Pragmatic & Incongruent ones (M = 2.5). The main effect of Group was non-significant (F (3, 79) = 1.88, p = .14, partial η^2 = .07). The interaction of Meaning Type and Group was marginally significant (F (3, 79) = 2.49, p =.07, partial η^2 = .09. The difference between variables in the younger children group was markedly smaller than in the other groups, as seen in the following table:

Table 6.7: Interaction of Group and Meaning Type in comparison of Pragmatic & Incongruent (1) and Pragmatic & Congruent (2) Meanings

Group	Meaning	*M*	*SE*
ASD	1 (P & I)	2.56	.16
	2 (P & C)	2.78	.17
Typical Adults	1	2.48	.13
	2	3.09	.13
Younger Children	1	2.7	.15
	2	2.86	.15
Older Children	1	2.27	.14
	2	2.56	.15

Discussion

The child study reported in this chapter was motivated, on the one hand, by a need to test whether a difference in performance between typical adults and adults with ASD reflected a developmental delay and, on the other hand, by a desire to extend existing research on scalar implicatures to other types of implicature. This discussion will treat these issues in turn.

Group Differences

The ratings data gave an interesting picture of participants' performance. Composite difference scores suggested that adults with ASD performed significantly worse than typical adults but not than typically developing children, and that typically developing children, in turn, performed significantly worse than typical adults. There were no significant differences in context sensitivity. ANOVAs on mean ratings, in contrast, suggested that the only significant difference was between typical adults and older children. Moreover, the analysis showed no significant difference in pragmatic and congruent versus pragmatic and incongruent meanings across the groups, suggesting a basic level of pragmatic competence in all groups. That is, all groups were capable of generating inferences according to context that would render felicitous an otherwise infelicitous utterance. These types of analysis are consistent: the first, the composite difference scores, finds a crucial difference between groups; the second, the ANOVAs, fails to find factors that explain this difference, suggesting that, in terms of the present analyses, adults with ASD are not behaving in a way that is qualitatively different from other participants. The results of these analyses are consistent, then, with a developmental delay in pragmatics in ASD.

The reaction-time data corroborate, somewhat, the evidence for a developmental delay in ASD. Without context, a pattern emerges across groups such that salient meanings are processed fastest, then nonsense, then pragmatic meanings; only the younger children deviate from this pattern. Moreover, adults with ASD were the slowest performers, and were significantly slower than the older children and typical adults. Likewise with contexts, adults with ASD are the slowest group of

participants, and are significantly slower than typical adults and older children. So far, the data are consistent with an ASD deficit and developmental delay. A potential inconsistency lies in the fact that whereas typical adults and younger children process congruent contexts fastest, then null contexts, then incongruent contexts, older children and adults with ASD process null contexts fastest, then congruent contexts, then incongruent contexts. We might reasonably expect a smoother developmental trajectory, with adults with ASD grouping, if with anyone, with the younger children, and the older children grouping with the typical adults. The actual developmental trajectory is jagged. One way to account for this jagged trajectory is by invoking the fact that older children were as fast as typical adults, and (non-significantly) faster than younger children; perhaps, then, the increase in the older children's speed of processing comes at the cost of an abnormal pattern among context types. If this explanation holds, then the reaction-time data as a whole are consistent with a developmental delay in pragmatics in ASD.

As we saw in the preceding chapter, the lack of differences in context-sensitivity scores should not necessarily be taken to imply that contextual processing is equivalent across the groups: that contexts affect processing to the same extent does not mean that they affect processing in the same way. To consider further how contexts affect processing, we turn to the models of utterance processing.

Models of utterance processing

The introduction to this chapter raised the issue of how to explain pragmatic abilities, and offered, as a possible explanation, models of utterance processing, on the one hand, and intelligence on the other. Regarding the models of utterance processing, the data offer limited evidence. Children's ratings speak to the models

as follows. All but three children replicated the basic pattern of preferring pragmatic meanings to nonsense and salient meanings to pragmatic meanings. Children gave higher ratings to salient meanings in context than to pragmatic meanings, which can be construed as evidence for the autonomous and graded-salience models of utterance processing and against the direct-access model. Children also gave higher ratings to salient meanings in incongruent contexts than to pragmatic meanings in incongruent contexts, which is possible evidence for autonomous and graded-salience models, and still rated salient meanings in incongruent contexts highly, which is possible evidence for the graded-salience model. This last datum, high ratings for salient meanings in incongruent contexts, can be considered only suggestive at most, however, for the following reason. Overall, contexts only appeared to affect ratings significantly in the 'salient + incongruent' condition. Perhaps, then, the apparent evidence for graded-salience is merely a sign that an ability to judge the appropriateness of an utterance to context is nascent. Because the reaction-time data show only one significant effect, they do not speak to the choice between models of utterance processing.

Implicature and intelligence

There was no evidence that IQ predicted performance on the pragmatic tasks.

Neither modelling of utterance processing nor IQ explains pragmatic performance in the current experiment; the respective differences among the groups await explanation. The previous chapter suggested an explanation in terms of pragmatic theory; that suggestion will be taken up in the following chapter, the General Discussion. This chapter will close by relating the current findings to those in the

literature: firstly, by comparing the findings with those of Surian et al (1996), on whose method the present design was based; secondly, by comparing the findings with those of the literature on scalar implicature.

Comparison with Surian et al

Given that this study was based on a critique of Surian et al (1996), we will find it instructive to consider how the findings in this study differ from those of Surian et al. Recall that Surian et al (1996) had children with autism, children with SLI and typically developing controls hear conversational extracts which were either straightforwardly acceptable or 'silly', silliness being mediated by violation of a Gricean maxim. Participants were matched by verbal mental age. The extracts were acted out with three puppets: a questioner and two answerers, one of which would give a silly response - in other words, a forced-choice format. The Test for the Reception of Grammar (Bishop, 1983) and a standard false-belief task were administered. The authors found that all – and only - children who passed the false-belief task correctly identified silly responses.

The tension between Surian et al's (1996) and this study stems from the performance of the typically developing samples in both studies; arguably the tension results from the markedly different methods used. Surian et al's 'silly' responses are equivalent to the present 'pragmatic meaning + null context' condition; whereas Surian et al's typical children rejected the silly meanings, the typical participants in the present study did not, but rather preferred them to nonsense. Given that Surian et al used a forced-choice format and the present study used a Likert-style scale, the difference in findings may not be crucial. The Likert-style scale used in the present study was easily mastered by children and allows

subtler responses and finer-grained comparisons between conditions. Perhaps the present study would have replicated Surian et al's data if it had used a forced-choice method, but in that case the study would have understated the children's pragmatic performance.

Perhaps, though, the tension reflects a genuine difference in performance, not just a methodological artefact. In that case, we can give the following explanation. Surian et al's typical group was considerably younger than the younger child cohort in the present study (6;7 and 7;5 respectively). Perhaps complete rejection of pragmatic meanings is a developmental stage that immediately precedes the stages suggested in this study. This difference may reflect progress towards an endpoint of a tolerance of apparent nonsense, from which pragmatic responses are sometimes derived. It is consistent with this view that there should be a preceding stage when all pragmatic meanings are rejected. To tease apart the two possibilities – a methodological artefact versus an earlier developmental stage – the Likert-style scale could be used with Surian et al's (1996) method or a forced-choice format with our materials. Since the forced-choice method could systematically underestimate children's pragmatic abilities, use of a Likert-style scale with Surian et al's (1996) method is preferable. Suggestive, here, is the fact that pilot testing with six-year-old children revealed intolerance of inferred responses. A more extensive data set would be needed for a proper analysis.

Relationship with literature

This study relates to the literature on scalar implicature in the following ways. Recall that Noveck (2001) found that, with scalar implicatures, 7-year-old children

were competent to some degree, 9-year-olds were significantly better, and adults were significantly better still. Other researchers, as the introduction showed, have found pragmatic abilities in younger children, even in 5-year-olds, by using acting-out rather than meta-linguistic tasks, minimising the number of distracters, and using only affirmative scales. The present study weakly supports Noveck's findings in that 7-year-old children showed nascent pragmatic abilities in distinguishing between pragmatic meanings and nonsense, and in that 9-year-olds were (non-significantly) better than 7-year-olds and worse than typical adults. The weakness of this support should be stressed, however. This study did not include younger children, a decision that was made because pilot testing with younger children (from 4-year-olds to 6-year-olds) revealed that they found the task too demanding and frustrating. This striking difference in abilities in younger children raises the question of whether the relatively low abilities found in 7-year-olds and 9-year-olds in this study are caused by the method, that is, by having children make a meta-linguistic judgement rather than acting out their response. But, as the introduction argued, particularised conversational implicatures do not lend themselves to acting-out strategies. Granted that we could devise experiments to make particularised non-scalar implicatures constrained enough to be equivalent to scalar implicatures; such constrained implicatures would arguably be ecologically invalid in virtue of their being so constrained. Given such problems it is understandable that developmental and clinical studies would focus on scalar implicatures; but if we want to provide empirical evidence as well as theoretical arguments for the equivalence of scalar implicatures and particularised non-scalar implicatures, then perhaps meta-linguistic tasks, such as the one used in this study, are indispensable.

Future studies could, nevertheless, work on reducing task demands as much as possible.

We turn, now, to a general discussion.

Chapter 7: General Discussion

Summary of findings

This thesis' ultimate aim was to investigate the common claim that autism impairs pragmatics. In particular its aim was to ask, in light of the different levels of pragmatics, which types of pragmatics are impaired by AD/ASD. The introduction to the thesis outlined the following possible levels of pragmatic impairment:

(1) Difficulties in assigning semantic values to indexical and otherwise context-sensitive expressions, and to ambiguous expressions: for example, deictic expressions such as pronouns or tensed expressions, and lexical ambiguity.
(2) Difficulties in understanding contextually calibrated lexical items: for example, *ad hoc* concepts such as context-sensitive uses of adjectives (e.g. 'fresh').
(3) Difficulties in saturating and freely enriching minimal propositions
(4) Difficulties in understanding implicatures.

The thesis used this framework of pragmatics to explore the existing literature, finding the best evidence for level (1) impairment and hints of level (2) and level (3) impairment, but an absence of evidence for level (4) impairment. The thesis also used the framework to explored pragmatics naturalistically in a corpus of natural-language data, finding again the clearest evidence for level (1) impairment and hints of level (2) and level (3) impairment, but again an absence of evidence for level (4) impairment. The naturalistic study also suggested a number of possible explanations of pragmatic phenomena, including ad hoc concept construction and cognitive modelling of utterance processing.

The literature review and the naturalistic study both highlighted a lack of evidence for level (4) impairment; given the centrality of level (4) phenomena to pragmatics, it is a crucial level to investigate in any supposedly pragmatic disorder. Accordingly the thesis explored pragmatics experimentally, focusing on level (4) pragmatics by testing participants' understanding of particularised conversational implicatures in null, congruent and incongruent contexts. Analyses using a summary statistic, a composite difference score, suggested that adults with AD deviated most from the ideal pattern of responding, followed by younger children, followed by older children, followed by typical adults; the typical adults deviated from the ideal significantly less than any other group. A second summary statistic, context sensitivity, showed no systematic differences between groups. These analyses were corroborated by a comparison of pragmatic and congruent and pragmatic and incongruent meanings across the groups, which found a basic pragmatic competence in all groups. Analyses using ANOVAs did not identify a factor to explain the differences identified by the composite difference scores. The data were, nevertheless, consistent with a developmental delay in AD. Reaction-time data broadly supported a developmental delay, too, with adults with AD being the slowest performers, and significantly slower than both older children and typical adults. A grouping of adults with AD with the older children in terms of processing null contexts fastest, then congruent contexts, then incongruent contexts was interpreted as implying that older children sacrifice normal processing to speed.

Attempts to test models of utterance processing on these data met with varying degrees of success. The models proved useful in characterising the typical-adult data. The data did not fit well with an autonomous model of utterance processing; it

fit reasonably well with a thoroughly interactive or ‘direct-access’ model; but it fit best the compromise graded-salience model. The models did not prove useful with the remaining groups, however, because, owing to substantial variance, there were insufficient points of clear comparison to fit any model of utterance processing to the data. It is important to ask why the attempt was unsuccessful in fitting particular models of the utterance processing. The various models cross-sectioned the logical space for models of utterance processing, covering serial, interactive and parallel possibilities; while other models are surely possible they would just as surely overlap, so that we would reasonably have expected some evidence to support one of the three possibilities. A likely answer is that the models cannot cope with such high level pragmatics as particularised conversational implicatures. Modellers of utterance processing have cut their teeth on lexical ambiguity (when there are literal and figurative alternatives), phrasal idioms, and metaphors, phenomena that, with the possible exception of metaphors, sit lower in the levels of pragmatics than do implicatures. Any modeller should surely expect that higher-level pragmatics will, in virtue of more complex processing demands, encourage more variation in dependent measures such as ratings and reaction times, and this expectation is still greater with young children and clinical-pragmatic groups, populations whose demonstrable pragmatic skills are more marginal. However, a modeller would rightly want to account for higher-level pragmatics since it is crucial to utterance processing. Future studies could test the models in different populations on more constrained but still higher-level pragmatics such as scalar implicatures. Finally, an attempt to test whether IQ predicted children’s performance on the pragmatic task found no significant effect. To a considerable extent, then, although the present

studies have found interesting differences between groups, they have not found a clear way to explain the differences.

Possible Explanations

It has been a recurring theme in the preceding two chapters that equivalent context sensitivity does not mean that all groups are doing the same things with contextual information. The current study tried unsuccessfully to apply models of utterance processing to the clinical and developmental data to probe what is done with contextual information. Future studies could profit from borrowing theories from linguistic pragmatics. To illustrate how such explanations might run, we will consider how one major pragmatic theory, Relevance Theory, might be usefully tested with clinical and developmental groups

Relevance Theory was discussed in some detail in the chapter on the typical adult study. Readers will recall that relevance theorists believe in communication, and indeed cognition, that is geared towards a search for relevance, where relevance is a function of maximised cognitive effects and minimised processing effort. The theory has been applied extensively to naturalistic data with good results (Huang, 2007). Increasingly, the theory is also tested experimentally using two different approaches. On the first approach, the researcher derives distinctive predictions from Relevance Theory about a specific phenomenon and tests these experimentally. Breheny, Katsos and Williams (2006), for instance, focused on generalised conversational implicatures, such as, ‘The shoes are black or brown’: the disjunction is logically true if the shoes are just black, the shoes are just brown, or the shoes are both black and brown; but there is a pragmatic implication that the

shoes are black or brown but not both. Whereas classical theorists took this implication to arise by default, relevance theorists take it to arise contextually. Breheny et al (2006) tested generalised conversational implicatures in neutral and inappropriate contexts, finding that the pragmatic implication did not arise in such contexts, casting doubt on its supposedly default nature. On the second approach, the researcher derives predictions from the relevance principles themselves, manipulating cognitive effects and processing effort. Van der Henst and Sperber (2004), for instance, manipulated these properties using a relational-judgement task, the Wason selection task, and a simple time-telling task, finding results in line with relevance-theoretic principles.

Relevance Theory has already featured in clinical research. Happé (1993) notably tested relevance-theoretic predictions about similes, metaphors and ironies with an autistic population. But it was not the relevance principles themselves but the involvement of theory of mind that was tested. Future studies could look for a more direct link to relevance principles by testing the extent to which different groups of participants can engage in relevance-based processing. Experiments could be so structured that possible explanatory factors could be used to predict scores on tests of implicature, metaphor and irony comprehension. For instance, the Wason Selection Task, which has been used in simplified versions with children as young as 3-years old (Cummins, 1996), could be used to score relevance-based processing, alongside an advanced theory-of-mind test such as the Strange Stories Task (Happé, 1994). If systematic differences between groups were found, an approach such as this, drawing on prominent linguistic theories, could provide valuable alternative

explanations of autistic language deficits to those limited explanations seen in the big three theories of autism as described in the naturalistic chapter.

Methodology

The experimental method used in this thesis has certain advantages; in particular the experimental materials and format have advantages over those of Surian et al (1996). The experiments in this thesis focused, not on the enigma of maxim violation, but on maxim flouting in context, and thereby gained surer access to pragmatics than did Surian et al. Moreover the experiments used a Likert-style scale that was easily learnt by children and that allowed subtler comparisons between conditions than Surian et al's forced choice format. And, finally, the experiments were tied to the models of utterance processing as an intended alternative explanation to theory of mind.

But the method also has certain disadvantages that would need to be remedied before any future studies used this format. Variance in participants' performance meant that, apart from the study on typical adults, the studies did not coherently support particular models of utterance processing, and hence participants' behaviour lacks a firm explanation. To counteract this defect, future studies could use tighter materials and more constrained phenomena such as scalar implicatures; materials could further be improved by presenting each scenario in all conditions to minimise the effects of typicality, lexical familiarity and so on. While the summary statistics provided a useful way of teasing underlying trends out of variance, their use could be modified. These statistics would benefit from the following simplification:

instead of a measure against an idealised norm, the developmental and clinical groups could more simply be measured directly against the typical adult means; this adaptation would arguably also be more objective. Future studies could also use an alternative dependent measure. For, although the Likert-style scale was easily learnt by all participants, it admits variance as follows. The scale called on meta-linguistic insight in the sense that participants had to judge the language used for its meaningfulness. Thus the scale was sensitive to variance in participants' linguistic skills and in their insight into their (and others') linguistic skills.

It is worth dwelling on the question of whether meta-linguistic tasks must be avoided; for such tasks, when used guardedly, are arguably essential to pragmatic research. The term 'meta-linguistic' has at least two senses that are relevant for present purposes; the present experiments were arguably meta-linguistic in both senses. Recall that participants were asked to give a rating as a function of how good a response was, how natural it sounded, how much sense it made. This task could be interpreted as a kind of stylistic judgement. But the task could also be interpreted as being akin to a truth-value judgement, in that participants might make a judgement on the proposition expressed by the language. Future studies would do well to isolate these meta-linguistic components and, if necessary, avoid them. But, as previous chapters have argued, meta-linguistic tasks may be essential when assessing higher-level pragmatics, at least in the sense that a judgement is performed on an expressed proposition. For, although simpler acting-out strategies have unearthed pragmatic skills in children as young as five, who were previously thought to be pragmatically naïve, these strategies work most readily with constrained phenomena like presupposition and scalar implicature, rather than with

thoroughly particularised contexts. Using acting-out strategies in such contexts may be prohibitively complex. This drawback can be mitigated, however, by following Pouscoulous et al (2007) in carrying out complementary tasks that vary task demands by varying the meta-linguistic nature of a task.

At this point, the question naturally arises of whether the current method can be extended to further studies. Future studies are clearly desirable in this area, because unconstrained implicatures are crucial to pragmatics. However, any future study would profit from tightening the materials, from exploring more constrained phenomena with modelling, and from minimising meta-linguistic demands in at least some tasks. With these adaptations, the existing method might be generalised.

Future Research

After the discussion of Relevance Theory as a possible explanatory framework in future studies, the preceding section has outlined various methodological motivations for future research. There are other, more conceptual motivations. The thesis has mentioned considerable variability in participants' performance; hence the use of summary statistics. This variability could be probed using other explanatory factors. For instance, tests of non-verbal intelligence could be extended to adult and ASD groups; tests of verbal intelligence could be added, alongside standardised tests of language skills such as syntactic and semantic ability and measures of literacy. And the variability among typical adults could be probed using behavioural games to identify distinct groups among participants perhaps correlating with personality-style traits among pragmatic hearers (Maria Ruz, *personal communication,* 2008).

Conclusion

This thesis has explored pragmatics in typical adults, typically developing children, and adults with autistic spectrum disorder, finding differences between the groups. Typical adults were more tolerant of pragmatics than were any other group; adults with ASD/AD most closely resembled the younger cohort of children. These data were interpreted as suggesting that there is autistic dysfunction in particularised conversational implicature, and hence in level (4) pragmatics. Methodological limits have been highlighted. The literature review and naturalistic study highlighted possible impairment to level (1) to (3) pragmatics. This study provides limited evidence for pragmatic dysfunction in autism at various different levels, and prompts future research taking account of the methodological limits.

Bibliography

Abbott, B. (2004). Definiteness and indefiniteness. In L.R. Horn & G. Ward (eds). *Handbook of Pragmatics* (pp. 122-149). Oxford: Blackwell.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: APA.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (5th ed.). Washington, DC: APA.

American Psychiatric Association. (2010). *APA DSM-V proposed revision: Asperger Disorder*. Retrieved from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=97#.

Braun, D. (2001). Indexicals. In *The Stanford Enclycopedia of Philosophy*. Retrieved from http://www.seop.leeds.ac.uk/entries/indexicals/

Baltaxe, C.A.M. (1977). Pragmatic deficits in the language of autistic adolescents. *Journal of Pediatric Psychology 2* (4), 176-180.

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: cognitive or affective? *Journal of Autism and Developmental Disorders, 18*, 379-402.

Baron-Cohen, S. (1989). Perceptual role-taking and proto-declarative pointing in autism. *British Journal of Developmental Psychology, 7*, 113-127.

Baron-Cohen (1995). *Mindblindess: An essay on autism and theory of mind.* Cambridge, Ma: MIT Press/Bradford Books.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a 'theory of mind'?. *Cognition, 21*, 37-46.

Barsalou, L.W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A.C. Collins, S.E. Gathercole, & M.A. Conway (Eds.) *Theories of memory*, 29-101. London: Lawrence Erlbaum Associates.

Bishop, D. V. M. (1983). Comprehension of English syntax by profoundly deaf children. *Journal of Child Psychology and Psychiatry, 24*, 415-434.

Bishop, D.V.M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communication impairment in children. *Journal of Child Psychology and Psychiatry,* 39, 879-891.

Bishop, D.V.M, & Adams, C. (1992). Comprehension problems in children with specific language impairment: literal and inferential meaning. *Journal of Speech and Hearing Research, 35*, 119-129.

Bishop, D.V.M., Chan, J., Adams, C., Hartley, J., & Weir, F. (2000). Conversational responsiveness in specific language impairment: Evidence of disproportionate pragmatic difficulties in a subset of children. *Developmental and Psychopathology, 12*, 177-199.

Bishop, D.V.M, & Norbury, C.F. (2005). Executive function in children with communication impairments, in relation to autistic symptomatology: I. Generativity. *Autism, 9*, 7-27.

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

Braine, M.D.S., & Rumain, B. (1981). Development of comprehension of 'or': evidence of a series of competencies. *Journal of Experimental Child Psychology, 31*, 46-70.

Breheny, R. (2006). Communication and folk psychology. *Mind and Language, 21* (1), 74-107.

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

Chierchia, G. Fox, D., & Spector, B. (*forthcoming*). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Malenborn & K. von Heusinger (Eds). *Handbook of Semantics*, Mouton de Gruyter

Clark, H.H. (1996). *Using Language*. Cambridge: CUP.

Crul, T.A.M., & Peters, H.F.M. (1976). *Auditieve discriminatietest (Auditory discrimination test)*. Lisse: Swets & Zeitlinger.

Cummings, L. (2007). Clinical pragmatics: a field in search of phenomena? *Language and Communication, 27* (4), 396-432.

Cummings, L. (2008). *Clinical pragmatics.* Cambridge: CUP.

De Villiers, J., Stainton, R., & Szatmari, P. (2007). Autism spectrum disorders and the semantics/pragmatics boundary: A case study in philosophy and the empirical. *Midwest Studies in Philosophy, 30* (1), 292-317.

De Villiers, J., Myers, B., & Stainton, R. (*forthcoming*). Differential pragmatic abilities in ASD. In M. Maccaulay & P. Garces (Eds). *Context and Pragmatics*. Antares: Toronto.

Diamond, A., Prevor, M., Callender, G., & Druin, D.P. (1997). Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. *Monographs of the Society of Research in Child Development, 62,* 1-207.

Eisenmajer, R., & Prior, M. (1991). Cognitive and linguistic correlates of theory of mind in autistic children. *British Journal of Developmental Psychology, 9*, 351-364.

Fine, J., Bartolucci, G., Szatmari, P., & Ginsberg, G. (1994). Cohenisve discourse in pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24* (3)*,* 315-329.

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology, 1*, 329-342.

Gernsbacher, M. A., & Robertson, R.R.W. (2001). The definite article *the* as a cue to map thematic information. In M. M. Louwerse & W. Van Peer *Thematics: Interdisciplinary studies.* Amsterdam/Philadelphia: John Benjamins.

Gibbs, R. W. Jr. (1994). *The poetics of mind: Figurative thought, language and understanding.* New York: Cambridge University Press.

Gibbs, R.W. Jr. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics, 34* (4), 457-486.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 7* (1), 183-206.

Giora, R. (2002). Literal vs figurative language: Different or equal? *Journal of Pragmatics, 34*, 487-506.

Giora, R., & Fein, O. (1999a). On understanding familiar and less familiar figurative language. *Journal of Pragmatics, 31*, 1601-1618.

Giora, R., & Fein, O. (1999b). Irony: Context and salience. *Metaphor and Symbol, 14* (4), 241-257.

Graetz, P, De Belser, R., & Willmes, K. (1992). Aachen Aphasia Test. Lisse: Swets & Zeitlinger.

Grant, J., Karmiloff-Smith, A., Gathercole, S., Paterson, S., Howlin, P., Davies, M., & Udwin, O.(1997). Verbal short-term memory and its relation to language acquisition in Williams Syndrome. *Cognitive Neuropsychiatry, 2* (2), 81-99.

Grant, J., Valian, V., & Karmiloff-Smith, A. (2002). A study of relative clauses in Williams syndrome. *Journal of Child Language, 29*, 403-416.

Gregoromichelaki, E., Kempson, R., Purver, M., Mills, G., & Cann, R. (*forthcoming*). Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse.*

Grice, H.P. (1969/1989). Utterers meanings and intentions, The Philosophical Review, 68. Reprinted in H.P. Grice (1989). *Studies in the Way of Words*. Cambridge, Ma: Harvard University Press.

Grice, H.P. (1975). Logic and Conversation. In D. Davidson & G. Harman (Eds). *The logic of grammar*, Encino, CA: Dickenson, 64-75.

Hala, S., Pexman, P.M., & Glenwright, M. (2007). Priming the meaning of homographs in typically developing children and children with autism. *Journal of Autism and Developmental Disorders, 37*, 329-340.

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

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

Happé, F.G.E. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry, 37* (7), 873-877.

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

Healey, P (1997). Expertise or expert-ese: The emergence of task-oriented sub-languages. In M.D. Shafto & P. Langley (Eds). *Proceedings of the 19th Annual Conference of the Cognitive Science Society*, 301-206.

Hermelin, B., & O'Connor, N. (1967). Remembering of words by psychotic and subnormal children. *British Journal of Psychology, 58*, 213-218.

Hobson, R.P. (1990). On the origins of self and the case of autism. *Development and Psychopathology, 2*, 163-181.

Honeck, P.R., Welge, J., & Temple, J.G. (1998). The symmetry control in tests of the standard pragmatic model: The case of proverb comprehension. *Metaphor and Symbol, 13*, 257-273.

Huang, Y. (2007). *Pragmatics*. Oxford: OUP

Hughes, C., & Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object: Its implications for theories of autism. *Developmental Psychology*, *29*, 498–510.

Jolliffe, T., & Baron-Cohen, S. (1999). The strange stories test: A replication with high-functioning adults autism or Asperger syndrome. *Journal of Autism and Developmental Disorders, 29*, 395-404.

Joseph, R.M., McGrath, L.M., & Tager-Flusberg, H. (2005). Executive dysfunction and its relation to language ability in verbal school-age children with autism. *Developmental Neuropsychology, 27* (3), 361-78.

Karmiloff-Smith, A., Klima, E., Bellugi, U., Grant, J., & Baron-Cohen, S. (1995) Is there a social module? Language, face processing and theory of mind in subjects with Williams syndrome. *Journal of Cognitive Neuroscience, 7* (2), 196-208.

Katz, A.N., & Ferretti, T. (2001). Moment-by-moment reading of proverb in literal and non-literal contexts. *Metaphor and Symbol, 16*, 193-221.

Kempson, R., Meyer-Viol, W., & Gabbay, D. (2001). *Dynamic Syntax: The Flow of Language Understanding*. Oxford: Blackwell.

Kjelgaard, M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Lnaguage and Cognitive Processes, 16*, 287-308.

Langdon, R., Davies, M. & Coltheart, M. (2002). Understanding minds and understanding communicated meanings in schizophrenia. *Mind & Language, 17*, 68-104.

Laws, G., & Bishop, D. (2004). Pragmatic language impairment and social deficits in Williams syndrome: A comparison with Down's syndrome and specific language impairment. *International Journal of Language Communication Disorders, 39*, 45-64

Leslie, A.M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology, 6*, 315-324.

Leslie, A. M., & Roth, D. (1993). What autism teaches us about metarepresentation. In S. Baron-Cohen, H. Tager-Flusberg & D.J.Cohen. *Understanding other minds*. Oxford: OUP.

Levinson, S.C. (1983). *Pragmatics.* Cambridge: CUP.

Levinson, S.C. (2000). *Presumptive meanings: The theory of generalised conversational implicature.* Cambridge, Ma: MIT Press.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24* (5), 659-685.

Martin, I., & McDonald, S. (2003). Weak central coherence, no theory of mind, or executive dysfunction? Solving the problem of pragmatic language disorders. *Brain and Language, 85*, 451-466.

Mills, G. (2007). *Semantic co-ordination in dialogue: the role of indirect interaction*. PhD thesis, Queen Mary University of London.

Minshew, N.J., Meyer, J., & Goldstein, G. (2002). Abstract reasoning in autism: A dissociation between concept formation and concept identification. *Neuropsychology, 16*, 327-334.

Mitchell, R.L.C., & Crow, T.J. (2005). Right hemisphere language functions and schizophrenia: The forgotten hemisphere. *Brain, 128*, 963-978.

Murphy, G. L. (1997). Polysemy and the creation of new word meanings. In T. B. Ward, S.M. Smith & J. Vaid (Eds.). *Creative thought: An investigation of conceptual structures and processes*, 235-265. Washington, DC: American Psychological Association.

Musolino and Lidz, J. (2006). Why children are not universally successful with quantification. *Linguistics, 44* (4), 817-852.

Musolino J., Crain S. and Thornton R. (2000). Navigating negative quantificational space. *Linguistics, 38* (1),1-32.

Norbury, C.F. (2004). Factors supporting idiom comprehension in children with communication disorders. *Journal of Speech, Language and Hearing Research, 47*, 1179-1193.

Norbury, C.F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology, 90*, 142-171.

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

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

Noveck, I.A., Bianco, M., & Castry, A. (2001). The costs and benefits of metaphor. *Metaphor & Symbol, 16* (1&2), 109-121.

Noveck, I.A., Guelminger, R., Georgieff, N., & Labruyere, N. (2007). What autism can tell us about 'Every…not sentences.' *Journal of Semantics, 24* (1), 73-90.

Nunberg, G., Wasow, T., & Sag, I.A. (1994). Idioms. *Language, 70* (3), 491-538.

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

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures. *Language Acquisition, 12*, 71-82.

Peleg, O., Giora, R., & Fein, O. (2004). Contextual strength: The whens and hows of contextual effects. In I. Noveck & D. Sperber (Eds.), *Experimental Pragmatics*, 172-186. Basingstoke: Palgrave.

Perkins, M. R., Dobbinson, S., Boucher, J., Bol. S, & Bloom, P. (2006). Lexical knowledge and lexical use in autism. *Journal of Autism & Developmental Disorders, 36*, 795-805.

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 & Developmental Disorders*, *39*, 607-618.

Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). Processing costs and implicature development. *Language Acquisition, 14* (4), 347-375.

Recanati, F. (2002). Unarticulated constituents. *Linguistics and Philosophy, 25*, 299-345.

Recanati (2003). *Literal Meaning*. Cambridge: CUP.

Rinehart, N.J., Bradshaw, J.L., Moss, S.A., Brereton, A.V., & Tonge, B.J. (2000). Atypical interference of local detail on global processing in high functioning autism and Asserger's disorder. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 41* (6), 769-778.

Roberts, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology*, *12*, 105–126.

Ropar, D., & Mitchell, P. (1999). Are individuals with autism and Asperger's Syndrome susceptible to visual illusions? *Journal of Child Psychology, Psychiatry and Allied Disciplines, 40* (8), 1238– 1293.

Schiffer, S.R. (1972). *Meaning*. Oxford: OUP.

Smith, C. L. (1980). Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, *30*, 191-205.

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

Stainton, R.J. (2004). The pragmatics of non-sentences. In L. Horn & G. Ward (Eds.), *The Handbook of Pragmatics*, 266-287.

Strawson, P.F. (1964). Intention and convention in speech acts. *The Philosophical Review, 73* (4), 439-460.

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

Tager-Flusberg, H. (1993). What language reveals about the understanding of minds in children with autism. In S. Baron-Cohen, H. Tager-Flusberg & D.J. Cohen, *Understanding other minds: Perspectives from autism*. Oxford: OUP.

Tager-Flusberg, H., Calkins, S., Nolin, T., Baumberger, T. Anderson, M, & Chadwick-Dias, A. (1990). A longitudinal study of language acquisition in autistic and Down syndrome children. *Journal of Autism and Developmental Disorders, 20*, 1-21

Tager-Flusberg, H., & Joseph, R.M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transaction of the Royal Society of London, B, 358*, 303-314.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, Ma: MIT Press.

Turner, N.E., & Katz, A.N. (1997). Evidence for the availability of conventional and of literal meaning during the comprehension of proverbs. *Pragmatics and Cognition, 5*, 203-237.

Van der Henst, J.B., & Sperber, D. (2004). Testing the cognitive and communicative principles of relevance. In I. Noveck & D. Sperber (Eds.), *Experimental Pragmatics*, 172-186. Basingstoke: Palgrave.

Van der Lely, H.K.J., & Christian, V. (2000). Lexical word formation in grammatical SLI children: A grammar-specific or input-processing deficit? *Cognition, 75*, 33-63.

Wassenaar, M.E.D., & Hagoort, P. (1994). *The offline observation instrument for grammatical comprehension*. Unpublished manuscript.

Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In N. Burton Roberts (ed.), *Advances in Pragmatics*, 230-260. Basingstoke: Palgrave.

Wilson, D., & Sperber, D. (2002). Relevance theory: A tutorial. In Y. Otsu (Ed.), *Proceedings of the Third Tokyo Conference on Psycholinguistics*, 45-70. Tokyo: Hituzi Shobo.

Winer, G.A., Cottrell, J.E., Mott, T., Cohen, M., & Fournier, J. (2001). Are children more accurate than adults? Spontaneous use of metaphor by children and adults. *Journal of Psycholinguistic Research, 30* (5), 485-496.

Appendix: Streamlined Corpus

KEY

Deixis

T	*Tense*
As	*Aspect*
P	*Deictic pronoun*

Other Pragmatic

D	*Disambiguation*
I	*Idiomatic Expression*
If	*Informativeness*
S	*Simile*
H	*Hyperbole*

Semantic

E	*Ellipsis*
Md	*Modality*
N	*Negation*
Q	*Quantification*
Re	*Reference*
Ls	*Semantic lexical access*

Structural

A	*Use of in-/definite article*
M	*Inflectional morphology*
Lp	*Phonological lexical access*
Pr	*Preposition*
Rf	*Reflexive pronoun*

Context	Anomaly	Code	Diagnosis
I rode on a horse but I wasn't standing up on it.	I will fall off. I don' think it will get me.	T; P	PLI
	You know Zorro in the film. He can…it's true..		PLI
	It can stand on three horse and control them..	P; M	
	at the same time and tell them which way	E	
	Them are a bit dangerous	P	
	He was…erm…the headmistress of the school…	P	PLI
	driving the bus		
	Them just sink	P	HFA
Have you ever ridden on a horse?	Yes. It's always itchy now.	P	HFA

	A man fainted on the car and a man tried …		SLI
	to blow air out of his mouth and he couldn't..	P	
	do it and he were dead.	M	
Do you ever take your dogs out for a walk?	Yeah, sometimes I do. Sometimes they like…		SLI
	to run and sometimes like to scratch theirself.	Rf	
	It could be Fred (Fred not introduced; name given by a child to a figure in test materials)	Re	SLI
No antecedent	It's very dangerous.	P	PLI
	Me want to do this	P	HFA
	[Child always referred to self as 'Mine']	P	HFA
Do they slip up and down when you walk?	No, they slip down and up	I	HFA
You've got a little while left, Gina.	Gina want a thousand whiles left. Gina like the whiles at here.	M; I	HFA
Do you get poorly quite a lot or are you quite healthy?	I don't get poorly quite a lot	I/N	PLI
What's happening there?	The man's doing his temperature	I	SLI
What was wrong with you?	I had sick and a bad back	I	SLI
	I don't see how you can spend time, and…	I	HFA
	I certainly don't see how you could…		
	enjoy it, because spending time is not…		
	something you can do.		
	You can only actually spend money.		
When are you taking off for Brisbane?	We're not taking off, we're driving.	I	HFA
We started recylcling but we ran out of enthusiasm. It's got flies on it.	You could catch a disease	I	HFA
Hop on the scales	[Hops on the scales]	I	HFA
Cut it out, Daniel.	Repeats, 'Cut it out, Daniel'; cuts off point of pencil he's holding with a pair of scissors	I	HFA
	In the far away distance there's the sea	I	HFA
	I rode on a horse one time.		
What was it like?	Like [mimes]	I	HFA
Who do you think they might be?	People	If	PLI
Can you remember what our sandcastles are like?	Like a big castle	I/If	PLI
Do you know what had happened. To make your back go funny?	The doctor didn't know what had happened.	D	SLI
Can you stand to do some more?	[stands up]	D	HFA
I thought you told me you were hardy	No, I'm Jones.	D	HFA
Watch your hands!	[Stares at fingers]	D	HFA
	Robinson Crusoe was marooned on an island because he got half way across the road and he was too nervous to go the rest of the way.	D	HFA
How does that look?	By its eyes	D	HFA

Put your milk on the table.	[Pours milk on table]	D/E	HFA
Your mum's going to kill me when she sees your shirt.	No, she won't. She's a Christian.	D	HFA
Mind if I join you?	With this? [offers speaker adhesive tape]	D	HFA
Go and get yourself a drink from the toilet	[drinks from the toilet bowl]	D	HFA
Put your bottom on that carpet	[picks up a carpet tile, holds it to his bottom and walks off]	D	HFA
[Used inappropriately each night]	I'm going to bed uncharacteristically early	I	HFA
It was raining cats and dogs	[deep distress]	I	HFA
The wind will take your head off.	[fear]	I/H	HFA
Your room looks as if a bomb has hit it	Oh my God, has there been a bomb?	I/S	HFA
Have a drink on me.	[confusion]	I/S	HFA
It feels like I'm talking to a brick wall.	Daddy, why are you talking to a brick wall?	I/S	HFA
Go and sit down.	[sits down on spot]	I/D	HFA
Mother responds to child sticking his finger up at her with 'Naughty finger!'	[bites finger as though independently naughty]	I/D	HFA
What can you see?	Mum and boy	E	PLI
And what's this one?	Bag	E	PLI
Let's have a look at the next one.	With doctor	E	PLI
What's wrong with the boy?	Hurting	E	PLI
	There's a policeman [who] doesn''t look after him a lot.	E	PLI
What's happening there?	Temperature	E	PLI
Have you ever been to the seaside?	Wales	E	PLI
How did you build your sandcastle?	Sand	E	PLI
What's happening there?	Party	E	PLI
	A sandcastle [that] would reach the sky.	E	SLI
Where would you like to go?	To see my dad. [In] Wales.	E	PLI
I once went ice-skating when I was little but I fell over all the time.	Um. My mum fell over [when] she's skating.	E	PLI
	No. I [am] bigger than the other two.	E	PLI
What's happening in that one?	The boy [is] in hospital.	E	PLI
What do you think about that?	Dangerous	E	PLI
I've been to a rodeo.	Did you?	E	PLI
Can you tell me what's happening?	Cowboy	E	HFA
[No contextual help]	If you really hurt, you will.	E	HFA
	I [am] not allowed to play GAME.	E	HFA
	They may [have] gone to a field trip.	E	HFA
	I going batman. I already batman.	E	PLI
	He used to came to the nursery	M	PLI
Have you ever had chicken pox?	Yes, and it was Sunday and we usually go to Grandma's on Sunday	T/As	PLI
What have you had wrong with you?	Er, forgot now	T/As	PLI
Have you ever been really ill?	[squeaks]	T/If	SLI+aut

	I haven't		
	I haven't		
	For months		
	For a year and six days		
	I wish I'd looked after my teeth….		SLI+aut
	Otherwise I won't be able to talk.	T	
	When I was getting up, I used to fall…	T	SLI+aut
	back. I lie on the settee.		
	I have saw one of them on television	T/As	SLI+aut
	She runs away and hide	T/M	SLI+aut
Will you have a birthday cake at your party?	Two of them. Well mum get two of them. One for me. And there's another that.. chocolate on the bit..and we give it to the…put it in the party bag.	T	PLI
	He used to squirt me with his supersoaker. It was powerful. He can shoot it from down the road to that door.	T/As	PLI
Context: going past	They can get someone who's going to pass to help them.	T/As	PLI
When did you do that?	When we go on holiday	T	PLI
Have you ridden a horse?	My sister had	T	PLI
Is your mum having a baby?	No, but she will in August	A	HFA
	Otherwise he should fell in the mud.	T	HFA
	Then he's going to fell in the potty mud.	T; LS	HFA
	I didn't went to hospital.	T	HFA
	What happened? What do you hurt?	T	HFA
	That's what Brogan does with her brownies once.	T	PLI
	I draw some balloons	T/M	HFA
	And he dig a hole.	T/M	HFA
	[The] tide was in	A	PLI
	I've got [a] wrestilng ring and four fingers	A	PLI
	He's in [a] different class	A	PLI
	Took a picture where a man went flying. A man died.	A	PLI
	When we went to [the] seaside.	A	PLI
What do brownies do?	[A] kind of dance.	A	PLI
Who invited you to the party?	All the people	A/Q	PLI
How did you get there?	In the aeroplane.	A	PLI
Can you swim without … your armbands?	Not yet. She has to hold me on the tummy.	A	PLI
	They're going in sand.	Pr	PLI
And where are they?	In the beach	Pr	PLI
What's happening there?	The boys…climbing into all the rocks.	Pr	PLI
So when's your birthday?	At 2 February	Pr	PLI
	I've been to the holiday	Pr	PLI
And when you go to Las Vegas … where did you stay?	In a hotel at up the escalators.	Pr	PLI
	Mum and Ben. No, mum and boy	Lp	PLI

	A water cobe	Lp/Ls	PLI
	This is a man who was driving a racing car. He just fell onto another car.	Ls	PLI
	All they had to do is just look at a light inside my belly button.	Ls	HFA
	They may live in Brazil but nowhere else	Ls/Md	HFA
	Maybe in the Ilantic Ocean.	Lp	HFA
	I went with the peech. We do speech. We speak a lot.	Lp	HFA
	I think it was one of the ends. You know the ends of the day.	Ls	HFA
[For cow]	There's a moose over there.	Ls	PLI
	There are these people who stopped and said, 'Congratulations' and then drove off and phoned the AA>	Ls	HFA
We caught a train on holiday.	No, we caught a train between A and B… and another train between B and C… [and so on].	If	HFA

Appendix: Experimental Materials

Context **S = Salient, P = Pragmatic**	**Question**	**Answer** **N=Nonsense, S=Salient, P=Pragmatic**
S: Y hates tea with sugar	How would you like your tea today?	N: I'd like it with fries
P: Yesterday X spilt tea over Y		S: I'd like it without sugar
		P: In a cup this time (Quantity)
S: Y has a watch on	Could you please tell me the time?	N: It is a blue whale
P: The museum opens at ten		S: It is ten o'clock
		P: The museum's not open yet (Relation)
S: X works for the police	Does this bag here belong to you?	N: Blue cheese tastes very good
P: Y's friend Paul is very forgetful		S: That is not my bag
		P: Paul was sitting there before (Relation)
S: X doesn't really like milk	How would you like your coffee today?	N: With a bit of lettuce
P: X is shaking the jug		S: With a little sugar please
		P: Preferably in a mug, please (Quantity)
S: Y is fussy about food	How would you like your sandwich, Y?	N: I'd like it by post
P: just spilt milk everywhere		S: With a little lettuce, please
		P: Without any milk, thank you (Quantity)
S: Y has a sweet tooth	How would you like your chocolate cake?	N: Under that plant pot there
P: X just dropped his cake		S: With some cream on it
		P: In a bowl, thank you (Quantity)
S: Y loves scones with jam	How would you like your fruit scone?	N: I'd like it without war
P: X sneezed dropping his cake		S: I'd like it with jam
		P: How about on a plate?? (Quantity)
S: Y is really really fussy	How would you like your hair cut?	N: Cut it with butter please
P: X often uses hair clippers		S: I'd like it quite long

		P: With a pair of scissors (Quantity)
S: X cooks really, really badly P: Y is very hungry indeed	What shall we have for breakfast today?	N: A brick broken in two S: Sausage and beans on toast P: Eggs cooked in hot water (Quantity)
S: X cooks anything Y wants P: X is clueless about cooking	What food would you like to eat?	N: Nails in a wood plank S: Salmon steak poached in milk P: Fried eggs cooked with oil (Quantity)
S: Y really likes drinking coffee P: Y needs lots of energy	Shall I make us some coffee, Y?	N: Trees really grow very slowly S: Yes, I would like one P: Coffee would keep me awake (Relation)
S: X baked cakes for tea P: The dog likes chocolate cake	Has someone just eaten the chocolate cake?	N: Karen often wears a suit S: Yes, I have eaten it. P: The dog looks very happy (Relation)
S: X is writing a shopping list P: X only eats brown bread	What shall we have for lunch, Y?	N: A plate full of mud S: Cheddar cheese and grilled bacon P: Real bread and real cheese (Manner)
S: X likes his meat pink P: Y sometimes undercooks his food	Can you tell me what dinner is?	N: A desk made of wood S: A steak that's very pink P: It's a fully cooked sausage (Quantity)
S: X loves English cream tea P: X forgets to boil kettles	What are we having for tea this afternoon?	N: Green grass covered in mud S: Cake all covered in cream P: Tea made with hot water (Quantity)
S: X and Y just met P: Y likes to show off	Tell me, do you have any brothers?	N: I have about 400 pounds S: I have only two brothers P: I have about 500 brothers (Quality)
S: X is helping with Y's diet P: X is making a big cake	Y, do you have any more chocolate?	N: I have lots of marmite S: I have one more bar P: I have several tonnes more

		(Quality)
S: X is having a coffee	Do you have any more double cream?	N: I have some big potatoes
P: X is making party food		S: I have one more pot
		P: I have a whole lake (Quality)
S: X likes a few sausages	Do you have any sausages at all?	N: I have some chocolate muffins
P: X is making sausage pie		S: I have a few sausages
		P: I have miles of sausages (Quality)
S: X arranged that night's concert	Did you enjoy the concert this evening?	N: The plant is a rose.
P: The concert was very bad		S: This singing was very good
		P: It sounded quite like singing (Manner)
S: Ann really annoys most people	Say, X, don't you just hate Ann?	N: London really isn't in Spain
P: Ann's friend is behind X		S: I certainly don't like her.
		P: Isn't the weather lovely, X? (Relation)
S: X is organising Y's party	Do you have many friends, X?	N: I have 2 big hammers
P: X is shy but loves animals		S: I have about 20 friends
		P: I have three nice dogs (Relation)
S: X is Y's head master	Do you have many lessons at school?	N: I fish at the weekend.
P: Y works really really hard		S: I have 6 each day
		P: I am working every minute (Quality)
S: Y is on a diet	What snack would you like?	N: A glass with cracks in
P: X forgets things when baking		S: Some low fat biscuits please
		P: Cakes with sugar in them (Quantity)
S: Y needs some trainers for running	What do you want from the shoe-shop?	N: A tank of gold fish
P: Y can't make his mind up		S: I need some good trainers
		P: Just a pair of shoes (Quantity)
S: Y needs flowers for a birthday	What will you get from this florist?	N: Some dirty planks of wood
P: X's annoying Y with questions		S: I'll get some nice roses
		P: I think, probably, some flowers (Quantity)
S: Y's socks have many holes	Will you buy anything at the clothes shop?	N: Some fish-fingers and lemons
P: Y never plans what to buy		S: I need some new socks
		P: I just need some clothes

		(Quantity)
S: X is fussy about tea	Will you get anything at the tea shop?	N: I won't get any elephants
P: X is really boring Y		S: I'll get really good tea
		P: Yes. I'll get some tea (Quantity)
S: There is a new teacher	Do you like the new teacher?	N: The queen is not bald
P: X's testing the new teacher		S: Yes, he is very nice.
		P: Well, he doesn't smell bad. (Relation)
S: Y is very busy today	Shall we have a run?	N: Green beans are very healthy
P: Y just ate lots of food		S: Yes, in about an hour
		P: I would have to explode (Relation)
S: Y only really likes bananas	Will you buy something at the fruit stall?	N: I will have a reindeer
P: Y loves almost all fruit		S: A big bunch of bananas
		P: Yes, I think some fruit (Quantity)
S: X is buying Y's drink	Will you get something from the milkshake shop?	N: A pad of white paper
P: Y can never choose milkshakes		S: A nice smooth banana milkshake
		P: Any old milkshake will do (Quantity)
S: Y doesn't really like X	Why won't you play with me anymore?	N: I enjoy eating raw cabbage
P: X and John hate eachother		S: Because I don't want to
		P: Because I am John's friend (Relation)
S: X always eats with Y	Why won't you come for dinner tonight?	N: I'm the King of France
P: X bullies [Grace] very badly		S: I'm sorry, I'm ill today
		P: I'm good friends with Grace (Relation)
S: X and Y are friends	Hey, why aren't you talking to me?	N: I do the gardening daily
P: X and [] are enemies		S: I don't have time now
		P: Because I'm in [Harry's] gang (Relation)
S: X can be very mean	Why will you not be my friend?	N: I almost never eat cheese
P: X often hits people		S: I don't really like you
		P: My friends don't hit people (Relation)
S: Y doesn't really like Badger	Are you going to Badger's birthday party?	N: Sausages taste good with jam
P: Mole's going to the party		S: No, I am not going.

		P: I don't really like Mole (Relation)
S: X is holding a lead	Has someone walked the dog yet today?	N: Absolutely all dogs are mammals
P: The dog hates having walks		S: I took him just now
		P: He does look quite happy (Relation)
S: X needs to look smart	Do you like my new haircut, Y?	N: Turnips can be very muddy
P: X likes his hair fashionable		S: Yes, it looks very good.
		P: It is really fashionable, X (Relation)
S:Y has strong colour tastes	Why don't you paint your room blue?	N: I like sweet raspberry jam
P: Y gets really sea sick		S: Because blue is really horrid
		P: Because the sea is blue (Relation)