



Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders

Courtenay Frazier Norbury*

Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

Received 5 August 2004; revised 16 November 2004

Abstract

Lexical ambiguity resolution was investigated in 9- to 17-year-olds with language impairment (LI, $n = 20$), autistic spectrum disorder (ASD) *plus* language impairment (ALI, $n = 28$), ASD and verbal abilities within the normal range (ASO, $n = 20$), and typically developing children (TD, $n = 28$). Experiment 1 investigated knowledge of dominant and subordinate meanings of ambiguous words. The LI and ALI groups knew fewer subordinate meanings than did the ASO and TD groups. Experiment 2 used a modified version of the Gernsbacher, Varner, and Faust (1990) paradigm to investigate contextual facilitation and suppression of irrelevant meanings. All groups demonstrated contextual facilitation, responding quickly and more accurately to words following a biased context. However, children with ALI and LI did not use context as efficiently as did their peers without language deficit. Furthermore, for the LI and ALI groups, errors in the suppression condition reflected poor contextual processing. These findings challenge the assumptions of weak central coherence theory and demonstrate the need for stringent language controls in the study of autistic cognition. © 2004 Elsevier Inc. All rights reserved.

Keywords: Autism; Language impairment; Context; Ambiguity

* Fax: +1865 281 255.

E-mail address: courtenay.norbury@psy.ox.ac.uk.

Introduction

Successful language comprehension requires not only an understanding of words and utterances in isolation but also the ability to integrate utterances so as to build a rich, coherent mental representation of the objects and events specified in such utterances and the relations between them (Bishop, 1997). Specific language impairment (SLI) and autistic spectrum disorder (ASD) are two developmental disorders frequently characterized by poor comprehension of language (Bartak, Rutter, & Cox, 1975; Bishop, 1979, 1997; Kjelgaard & Tager-Flusberg, 2001). These comprehension deficits are particularly evident when individuals must integrate information within a context to infer implicit meanings (Bishop & Adams, 1992; Jolliffe & Baron-Cohen, 1999; Norbury & Bishop, 2002) or to resolve ambiguous expressions such as figures of speech and homographs (Happé, 1997; Kerbel & Grunwell, 1998; Norbury, 2004).

How are mental representations formed? Language comprehension is a dynamic process in which listeners are interpreting and integrating linguistic and nonlinguistic information from the very beginning of an utterance. The emerging representation of the utterance may be constrained by the semantic properties of individual words, the discourse context in which they occur, and ongoing syntactic analysis (Tyler & Marslen-Wilson, 1981). Furthermore, interpretation of a single utterance “facilitates” the understanding of subsequent utterances provided that the new information is relevant to the context. The structure-building framework of Gernsbacher (1990) provides a model for understanding this process. Gernsbacher proposed that as a representation is built up, memory cells will enhance activation of related meanings while at the same time suppressing activation of unrelated meanings. If unrelated meanings are not suppressed, a new representation will be initiated to accommodate this information. Some support for this framework can be found in computational models of sentence processing that have demonstrated that processing of sentence contexts involves activating appropriate information units while inhibiting (or suppressing) all other types of information (MacDonald, Pearlmutter, & Seidenberg, 1994).

Comprehension deficits in SLI and ASD

Children with SLI are likely to have difficulties during the early stages of contextual processing due to generally poorer vocabulary (for a review, see Bishop, 1997), weaker semantic representations of the vocabulary they do have (McGregor, Newman, Reilly, & Capone, 2002), and deficits in syntactic processing (Bishop, 1997). Studies of children’s discourse comprehension have further suggested that children with SLI experience deficits in contextual processing that may be out of proportion with their structural language abilities. For example, Bishop and Adams (1992) found that story comprehension scores for children with SLI were poor relative to those for peers even when sentence comprehension scores were partialled out in analysis of covariance. Furthermore, individuals with SLI were poor at both verbal and nonverbal (pictorially presented) stories, leading the authors to conclude that children with SLI have difficulty in integrating sequentially presented information.

An influential theory of the cognitive deficit in autism suggests that individuals with ASD also fail to process context efficiently but postulates that this is due to a cognitive difference in the normal drive for coherence (Frith, 1989; Happé, 1999). In other words, individuals with ASD process information in a piecemeal fashion rather than integrating information in a context to construct higher level meaning (Frith & Happé, 1994). For example, numerous studies have shown that individuals with ASD are less likely to provide appropriate pronunciations of homographs following supportive contexts [e.g., “Her dress/eye had a *tear* in it”] than are typically developing (TD) comparison group members (Happé, 1997; Jolliffe & Baron-Cohen, 1999; Lopez & Leekham, 2003). Similar findings have been reported across a number of different verbal tasks such as inferencing (Jolliffe & Baron-Cohen, 1999, 2000; Ozonoff & Miller, 1996) and resolving lexically and syntactically ambiguous sentences (Jolliffe & Baron-Cohen, 1999). Importantly, poor performance has been demonstrated in individuals with verbal IQ scores within the normal range (Jolliffe & Baron-Cohen, 1999). For such high-functioning individuals, comprehension deficits may arise from a specific inefficiency in integrating linguistic information in context.

However, recent investigations of language profiles of children with ASD have highlighted linguistic deficits in a substantial proportion that are similar to the linguistic profiles characteristic of SLI (Kjelgaard & Tager-Flusberg, 2001; Tager-Flusberg & Joseph, 2003). Furthermore, performance on the homograph task has been associated with general verbal ability. Snowling and Frith (1986) reported that children with autism who had good verbal skills performed as well as TD peers on the homograph task. This raises the possibility that individual differences in processing verbal context are related to individual differences in core language abilities (i.e., vocabulary and grammar) in both SLI and ASD populations (Norbury, 2004).

The role of inhibition in comprehension

Studies of poor comprehenders by Gernsbacher and colleagues (Gernsbacher & Faust, 1991; Gernsbacher & Robertson, 1999; Gernsbacher, Varner, & Faust, 1990) have suggested deficits not in the initial activation and integration of relevant information but rather in the suppression of irrelevant information. For instance, Gernsbacher et al. (1990) presented adult skilled and poor comprehenders with short sentences, half of which ended with an ambiguous word (e.g., “He dug with a *spade*”) and half of which ended with an unambiguous word of similar meaning (e.g., “He dug with a *shovel*”). Following this, a test word such as *ace* was presented, and participants were asked to judge whether the word was related to the sentence meaning. In both examples, *ace* should be rejected because it is unrelated to the meaning of either sentence. However, *ace* and *spade* do have associated meanings in a different context (i.e., playing cards). As comprehension of sentences proceeds incrementally, associated information may be activated as each word is encountered. Therefore, the playing cards meaning of *spade* may be initially activated. However, the structure-building framework suggests that if the activated information does not fit with the existing mental representation (built up around “digging”

and its associations), then this information will be rapidly suppressed. Comparing the time taken to reject contextually inappropriate test words such as *ace* following ambiguous and unambiguous sentence contexts provides a measure of how activated the inappropriate meaning of the ambiguous word was. College students showed initial interference and were slower to reject the test word following the ambiguous sentence. However, after a significant delay (850 ms), only those students with relatively poor language comprehension continued to show this effect (Gernsbacher et al., 1990). Persistent activation of irrelevant meanings prevents new information from being integrated into the existing representation, resulting in comprehension deficits for extended discourse.

Although Gernsbacher and colleagues have focused on adults with poor comprehension, recent evidence suggests that children with both SLI and ASD have considerable difficulties on tasks of verbal and nonverbal inhibition (Bishop & Norbury, 2005). However, the inhibition deficits were related to language ability rather than being a general feature of ASD (for a review of inhibitory function in ASD, see Hill, 2004). Therefore, it is possible that deficits in suppressing irrelevant information may be seen only in children experiencing language impairment (LI) regardless of autistic status.

In summary, research to date suggests that the ability to process meaningful context may fail for three distinct reasons: (a) a failure to access or activate relevant information, (b) appropriate activation but failure to integrate that information, and (c) appropriate activation and integration of relevant information but an inefficiency in suppressing irrelevant information. The current study was designed to investigate the source of comprehension problems observed in children with SLI and ASD.

The current study

The current study focused on lexical ambiguity resolution as a means of investigating individual differences in contextual processing. Ambiguous words such as *bank* are pervasive in the language, yet comprehension is rarely disrupted in the normal course of events because context readily suggests the correct interpretations of such words (Tabossi & Zardon, 1993). This suggests “selective access” to ambiguous meanings determined by contextual constraints (MacDonald et al., 1994). However, there is some evidence to suggest that the dominant, or more frequent, interpretations of ambiguous words are always at least partially activated (Swaab, Brown, & Hagoort, 2003; Syssau, Brouillet, & Groen, 2000). Context will, therefore, be more influential when the subordinate meaning of an ambiguous word is required (MacDonald et al., 1994).

Weak central coherence theory (Frith, 1989; Frith & Happé, 1994) predicts that “people with autism should show an inability to use context, for example, to disambiguate linguistic material” (Happé, 1997, p. 3). It also predicts that this deficit in extracting meaning from context should apply across all ages and abilities. However, this claim has been difficult to evaluate because most studies of contextual processing in autism recruit heterogeneous groups that vary enormously in verbal ability. Therefore, if the group as a whole performs poorly on a test of contextual process-

ing, it is unclear whether this is due to a core *linguistic* deficit or to a *cognitive* difference in integrating verbal information. The current study sought to tease these factors apart by comparing two groups of children with ASD with their TD peers and a nonautistic group of children with LI (Norbury, 2004, *in press*). The two autistic groups were segregated according to language status such that one group had linguistic impairments similar in kind and severity to the LI group, whereas the other group had scores on language assessments within the normal range. This provides an indication of whether deficits in contextual processing are specific to autism or are subject to more general linguistic influences.

Two experiments were conducted. Experiment 1 investigated children's knowledge of and access to both dominant and subordinate meanings of ambiguous words. Experiment 2 investigated contextual facilitation of relevant meanings and suppression of irrelevant meanings in sentence contexts. In general, the study sought to contrast two possible outcomes. First, if a failure to extract meaning from context is indicative of a cognitive difference associated with ASD, the two groups with ASD should perform more similarly to one another regardless of language status. Second, if deficits in contextual processing are related to core linguistic difficulties, the two groups with LI (as documented by standardized tests) should perform more similarly to one another regardless of autistic status.

Experiment 1: Words in isolation

Experiment 1 was designed to investigate children's knowledge of multiple meanings of ambiguous words. It is well recognized that inhibition of irrelevant meanings can occur only if both meanings are equally activated (Syssau, Brouillet, & Groen, 2000); therefore, Experiment 1 acts as a control task for Experiment 2. Children were asked to judge whether a picture matched a spoken ambiguous word. Pictures featured either the dominant or subordinate meanings of the words. Words were selected so that children would be familiar with both word meanings. Therefore, it was anticipated that there would be minimal differences in accuracy between dominant and subordinate meanings both within the groups and between the groups. However, it was possible that children with language impairment (both the LI and ALI group) would be less familiar with the subordinate meanings due to their generally poorer vocabularies. In terms of response time (RT), it was predicted that all groups would be slower to respond to subordinate items, assuming that dominant meanings were activated first.

Method

Participants

For this experiment, 74 children with communication impairments were recruited from specialist schools and units throughout southeast England. They were a subset of children taking part in studies by Norbury (2004, *in press*). Children who were

selected met the following criteria: between 9 and 17 years of age, nonverbal abilities within the normal range [standard score of 80 or above on the Performance subscale of the Wechsler Abbreviated Scales of Intelligence [WASI] (Wechsler, 1999)], monolingual English-speaking home environment, no sensorineural hearing impairment, and no evidence of neurological impairment. In addition, 28 TD children matched for age and nonverbal ability were selected. All children took part in both experiments.

Children were selected to represent a range of clinical diagnoses based on clinical report and documented medical diagnoses available in school files. They were grouped according to the procedure outlined in Norbury (2004). Briefly, children were divided on the basis of reported autistic features and scores on standardized tests of language functioning. According to clinical records, 49 children did have reported autistic features, whereas 25 did not. This was independently assessed using the Social Communication Questionnaire (SCQ) (Berument, Rutter, Lord, Pickles, & Bailey, 1999). This is a 40-item parental questionnaire that assesses all three aspects of the autistic triad: communication, social interaction, and restricted interests and behaviors. It is intended to be used as a screening instrument but provides good diagnostic agreement with the more extensive Autism Diagnostic Interview–Revised (ADI-R) (Berument et al., 1999; Bishop & Norbury, 2002; Lord, Rutter, & Le Couteur, 1994).

Language ability was assessed using three standardized measures: the British Picture Vocabulary Scales (BPVS 2nd edition) (Dunn, Dunn, & Whetton, 1997), which measures receptive vocabulary; the Concepts and Directions subtest of the Clinical Evaluation of Language Fundamentals (CELF-III^{UK}) (Semel, Wiig, & Secord, 2000), which measures understanding of increasingly complex sentences; and the Recalling Sentences subtest of the CELF-III^{UK}, which indexes expressive language ability. Children scoring below -1.25 *SD* on two of the three tests, or -2 *SD* on one test, were regarded as having LI. Using this analysis, children generally fell into one of three groups: LI ($n = 21$), ALI ($n = 29$), and ASD only (ASO, $n = 20$). A small number of children ($n = 4$) did not fit into any of these groups and were excluded from the study. Two children, one from the LI group and one from the ALI group, were later excluded for failing to complete all of the tasks, leaving a total sample of 68 clinical cases.

An additional group of TD children ($n = 28$) was recruited from a similar geographical area. Children in this group did not have a history of language or communication difficulties and scored within the normal range on all language measures.

Background measures

Table 1 reports the mean age, nonverbal ability scores, and scores on the selection measures for the three clinical groups and the TD group. The groups were matched for age, $F < 1$. Although all children were selected to have nonverbal abilities within the normal range, the LI group had significantly lower scores on the WASI Performance subscale than did the ASO and TD groups, $F(3, 87) = 3.38$, $p = .02$.

The results of the SCQ provide validation of clinician reports in that the two groups with reported autistic features (ALI and ASO) had significantly higher scores

Table 1

Mean scores (and standard deviations) on background and group selection measures for each group

<i>n</i>	LI 20	ALI 28	ASO 20	TD 28
Age (years)	12.59 (1.82)	13.03 (2.12)	13.07 (1.54)	12.46 (1.74)
WASI	96.20 ^a (9.90)	102.89 ^{ab} (11.73)	106.80 ^b (13.35)	107.14 ^b (10.70)
BPVS	83.15 ^a (11.00)	80.57 ^a (12.04)	104.00 ^b (14.08)	106.39 ^b (12.08)
Concepts	4.60 ^a (1.96)	4.86 ^a (1.63)	8.65 ^b (3.23)	11.57 ^c (2.13)
Recall	4.45 ^a (1.88)	4.43 ^a (1.77)	8.30 ^b (2.30)	9.82 ^b (1.72)
SCQ	12.63 ^a (7.78)	23.00 ^b (7.04)	26.32 ^b (5.15)	2.85 ^c (2.52)

Note. Standard deviations are in parentheses. Values with different superscripts are significant at $p < .05$. LI, language impairment; ALI, autistic spectrum disorder *plus* language impairment; ASO, autistic spectrum disorder and verbal abilities within the normal range; TD, typically developing; WASI, Performance subscale of the Wechsler Abbreviated Scales of Intelligence; BPVS, British Picture Vocabulary Scales; SCQ, Social Communication Questionnaire; Concepts, Concepts and Directions subtest of the Clinical Evaluation of Language Fundamentals (CELF-III^{UK}); Recall, Recalling Sentences subtest of the CELF-III^{UK}. Standard scores are reported for the WASI and BPVS, with a normal mean of 100 and a standard deviation of 15. Concepts and Recall standard scores have a normal mean of 10 and a standard deviation of 3. Total raw scores of SCQ are reported, with higher scores indicating greater severity of autistic symptoms. The cutoff score for pervasive development disorder is 15, whereas the cutoff score for core autism is 22.

on the SCQ than did the TD and LI groups, $F(3, 75) = 53.87$, $p < .001$. The LI group scores were also significantly different from those of the TD group, $p < .001$. It is not unusual for children with LI to exhibit raised thresholds of autistic-type behavior relative to TD children, even if the former do not meet full criteria for ASD (Bishop & Norbury, 2002). The SCQ provides cutoff scores for “diagnoses” of pervasive developmental disorder (PDD: scores of 15–21) and core autism (scores of 22 or above). Reassuringly, 95% of ASO children and 85% of ALI children had scores of 15 or above. Those that did not meet the cutoff for PDD tended to have borderline scores of 12–14 (except for one child in the ALI group who had a total score of 8). The profile of the LI group was less clear-cut. The majority of children had scores well below 15 (i.e., within the normal range). Approximately 30% of children scored within the PDD range, although most scores were borderline scores of 15. Three children scored in the autism range with scores above 21. It is important to realize that the SCQ taps behaviors that were reported to occur at a much earlier age of development (4–5 years), that none of the children in the LI group had ever had a diagnosis of PDD, and that there were no current professional or parental concerns about the presence of autistic behaviours. These data echo those from earlier studies reporting the heterogeneity of the LI population and the potential for change in behavior over time (Bishop & Norbury, 2002).

As expected, there were significant group differences on all language measures, BPVS: $F(3, 87) = 30.11$, $p < .001$; Concepts and Directions: $F(3, 87) = 49.77$, $p < .001$; Recalling Sentences: $F(3, 87) = 50.97$, $p < .001$. In all instances, the LI and ALI groups did not differ from one another and had significantly poorer scores than did both the ASO and TD groups. On the BPVS and the Recalling Sentences

subtest of the CELF-III^{UK}, the ASO and TD groups did not differ from one another. However, on the Concepts and Directions subtest, the ASO group had significantly lower scores than did the TD group, $p < .001$, even though their mean scores were within the normal range.

Materials

Words

A list of 62 ambiguous words such as *bank* was drawn up to include words that would be familiar to school-age children and for which both meanings of the words were nouns (e.g., *river bank* vs. *financial institution*). During pretesting, 30 10–11-year-olds completed a word association task in which the children were asked to write the first words that came to mind in response to the test words. In this way, dominant and subordinate meanings could be established. Seven words were highly polarized in that more than 80% of the children provided the dominant meanings in the word association task. Two words had balanced meanings. The complete list of words used in the study and the proportions of children responding with the designated meanings is provided in the Appendix.

Picture stimuli

Pictures representing the dominant and subordinate meanings of each word were downloaded from the Microsoft Clipart Design Gallery (<http://office.microsoft.com/clipart/default.aspx?lc=en-gb>) and were chosen to be as simple and unambiguous as possible. Ten adults were asked to name the complete picture set. Where there were consistent confusions (e.g., *palm tree* named as *island*), new pictures were selected and renamed by the adults. Pictures correctly named by at least eight of the adults were retained, leaving a total of 34 words, from the original list of 62, at this point. Finally, the task was pilot tested on 18 10- or 11-year-olds. After the test, 12 items were discarded because accuracy was less than 80% and/or time to respond was greater than 1100 ms. This left a total of 22 ambiguous words and 44 pictures in the current study.

Procedure

Ambiguous words were recorded directly onto a laptop computer by a male speaker of British English, and the experiment was run using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Children heard each word through headphones three times in the course of the task: once followed by a picture showing the dominant meaning, once followed by a picture of the subordinate meaning, and once followed by an unrelated picture. The word list was pseudo-randomized so that the same word did not appear in succession. Two lists were made with opposite orders of item presentation and were counterbalanced across participants.

Each child was tested individually in a quiet room at home or at school. Participants were asked to decide whether each picture could be one of the meanings of the

word and were instructed to press the “yes” or “no” button on the button box as quickly, but as accurately, as possible. At the beginning of each trial, a “Ready?” message appeared in the center of the screen. This remained until the experimenter pressed the space bar, ensuring that children were comfortable and attentive for each trial. Following each spoken word, there was a delay of 1 s followed by the target picture. The picture remained on the screen for a maximum of 5 s or until the child pressed the button box. RTs were recorded from the onset of each picture. No feedback on accuracy was provided. At the completion of each trial, the “Ready?” screen reappeared.

Results

Accuracy

Table 2 reports the mean number correct (out of 22 words) per group in each of the three conditions: dominant, subordinate, and unrelated. Error rates were low overall. A 4 (Group) \times 3 (Word Type: dominant, subordinate, or unrelated) analysis of variance (ANOVA) with repeated measures on the word type factor was calculated on the number correct. Because there were significant differences between the LI group and the ASO and TD groups on nonverbal intelligence quotient (NVIQ), all analyses were rerun with NVIQ scores (from the WASI Performance subscale) entered as a covariate. This did not alter the pattern of results, and the effect of the covariate was nonsignificant. Therefore, the original ANOVA results are reported. Post hoc analyses were conducted using the Games–Howell test for unequal variances. The effect size η^2 , which estimates the proportion of total variance accounted for by the independent variable, is also reported. This analysis yielded a main effect of type, $F(2, 184) = 43.13$, $p < .001$, $\eta^2 = .32$, $F(2, 42) = 17.73$,

Table 2
Mean numbers correct and mean RTs (in milliseconds) per group for Experiment 1: Words in isolation

		Dominant		Subordinate		Unrelated	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
LI	Score	19.90	2.00	17.10 ^a	3.90	20.65	1.14
	RT	820	245	907	218	1021	259
ALI	Score	19.89	1.97	17.37 ^a	3.56	21.07	1.61
	RT	805	202	928	220	957	171
ASO	Score	21.15	0.94	18.85 ^b	2.91	21.55	0.69
	RT	750	216	785	204	950	202
TD	Score	20.96	1.23	20.11 ^b	1.69	21.06	1.31
	RT	767	193	819	264	927	220

Note. Group means in the same column with different superscripts are significantly different at $p = .01$. The maximum score is 22. LI, language impairment; ALI, autistic spectrum disorder *plus* language impairment; ASO, autistic spectrum disorder and verbal abilities within the normal range; TD, typically developing; RT, response time.

$p < .001$, $\eta^2 = .46$, such that all groups were more accurate on unrelated meanings. A significant effect of group was also evident, $F(3, 92) = 7.53$, $p < .001$, $\eta^2 = .20$, $F(3, 63) = 16.01$, $p < .001$, $\eta^2 = .43$, with the LI and ALI groups obtaining significantly lower scores than the ASO and TD groups, which did not differ from one another, $ps = .01$. There was also a significant Group \times Word Type interaction, $F(6, 184) = 3.61$, $p = .002$, $\eta^2 = .11$, $F(6, 126) = 6.60$, $p < .001$, $\eta^2 = .24$. Tests of simple main effects showed that the groups did not differ significantly on accuracy of unrelated items, $F = 1.69$. There was an effect of group on dominant items, $F(2, 92) = 4.18$, $p = .008$, with post hoc analysis showing that the ALI group made marginally more errors than did the ASO and TD groups: ALI versus ASO, $p = .051$; ALI versus TD, $p = .08$. However, on subordinate meanings, the ASO and TD groups were significantly more accurate than the ALI and LI groups, $F(3, 92) = 6.46$, $p = .001$; LI versus ASO, $p = .03$, TD, $p = .007$; ALI versus ASO, $p = .04$, TD, $p = .007$.

Response time

Because the spread and variability in RTs was large, it was very difficult to detect true outliers, and the risk of eliminating extreme but valid RTs was great. Ulrich and Miller (1994) recommended that extreme RTs should not be removed because this may introduce considerable bias. Therefore, RTs were analyzed for all correct responses. Both mean and median RTs (which are likely to be more stable) were analyzed, but because these analyses produced the same pattern of results, only mean RTs are reported.

A repeated measures ANOVA revealed a significant main effect of word type, $F(2, 184) = 55.91$, $p < .001$, $\eta^2 = .38$, $F(2, 20) = 21.86$, $p < .001$, $\eta^2 = .51$, such that all groups were significantly slower to respond in the subordinate meaning condition. The main effect of group and the Group \times Word Type interaction were significant by items only: group, $F(3, 92) = 1.08$, $p > .10$, $F(3, 63) = 22.17$, $p < .001$, $\eta^2 = .51$; Group \times Word Type interaction, $F(6, 184) = 1.61$, $p > .10$, $F(6, 126) = 4.65$, $p < .001$, $\eta^2 = .18$. Individual items were not equated for frequency of occurrence or polarity, and this might make some items disproportionately difficult for some children. The lack of group effect by subjects stems from the large within-group variation on RTs.

Discussion

The low error rates suggest that, on the whole, both meanings of the ambiguous words chosen were known to the participants and the pictures of each meaning were clearly recognizable. However, the two groups with language impairment did make significantly more errors on subordinate items, suggesting that they might not be familiar with these less frequent, alternative word meanings. They also tended to respond more slowly to subordinate items but were not prone to make impulsive mistakes, as indicated by their performance on the unrelated items.

Experiment 2: Contextual facilitation and suppression

Experiment 1 suggested that, generally speaking, the children in this study were aware of both the dominant and subordinate meanings of ambiguous words but that the dominant words were more readily accessible. Experiment 2 investigated the same children's ability to process context so as to determine the correct meanings of ambiguous words in two conditions. The first condition considered contextual facilitation, that is, the ability to integrate information in a sentence to determine the most appropriate meaning. This was done by comparing accuracy and RTs to picture judgments following a neutral sentence (e.g., "He *ran* from the bank"—picture *money*) with those following a biased sentence context (e.g., "He *stole* from the bank"—picture *money*). The second condition explored suppression, that is, the ability to reject contextually irrelevant information. This was done by comparing accuracy and RTs to reject irrelevant ambiguous pictures following an ambiguous sentence (e.g., "He stole from the *bank*"—picture *river*) with those following a non-ambiguous sentence (e.g., "He stole from the *shop*"—picture *river*).

Gernsbacher and Faust (1991) did not specifically contrast dominant and subordinate meanings because they selected words with balanced meanings. The words in the current study varied in the extent to which meanings were polarized, so it was necessary to consider potential differences in dominant and subordinate meanings. There is evidence that the dominant meaning of an ambiguous word is always at least partly activated (Swaab et al., 2003). This suggests that contextual facilitation may be more pronounced for subordinate items. This may be particularly important for children with LI in the current study (LI and ALI) who demonstrated poorer knowledge of subordinate meanings out of context in Experiment 1. In terms of suppression, it has been demonstrated that when context constrains the subordinate meaning, both subordinate and dominant meanings may be active (Tabossi & Zardón, 1993). Therefore, interference effects may be more pronounced for subordinate meanings. Gernsbacher and Faust (1991) reported that skilled comprehenders did experience interference, but only immediately after sentence offset. After a delay of 850 ms, only poor comprehenders continued to experience interference. The extent to which children with poor language comprehension in the current study (LI and ALI) demonstrate interference will depend in part on the degree of contextual facilitation that they experience (Syssau et al., 2000).

Specific hypotheses: Contextual facilitation

Using a written presentation, Gernsbacher and Faust (1991) demonstrated that adults with poor comprehension experienced the same amount of contextual facilitation as did their more skilled counterparts. Using a spoken presentation, the first possibility is that all children in the current study will be faster and more accurate in response to biased contexts as opposed to neutral contexts.

A second possibility considers a specific cognitive deficit in extracting meaning from context that is thought to characterize individuals with autism (Happé, 1997). Following this framework, it is expected that the ALI and ASO groups will

show less contextual facilitation than will TD controls. This deficit is likely to be more evident in the understanding of subordinate items because these items may be more reliant on contextual information for disambiguation (MacDonald et al., 1994). No prediction is made about the performance of children with LI who do not show evidence of autism.

However, interpreting the biased context depends crucially on linguistic information (i.e., understanding the verb in the sentence and linking it to the ambiguous item, which is always the direct object). Therefore, children with LI may experience difficulty in using context efficiently. If this is the case, children in the LI and ALI groups might show less contextual facilitation than will TD controls and perhaps the ASO group.

Specific hypotheses: Suppression

Gernsbacher and colleagues have demonstrated that adults with poor comprehension are less able to suppress irrelevant meanings of ambiguous items than are their more skilled counterparts (Gernsbacher & Faust, 1991; Gernsbacher et al., 1990). Therefore, the first possibility is that the clinical groups with language comprehension deficits (LI and ALI) will show increased levels of interference relative to their TD peers, even after a delay of 1 s. TD children would not be expected to show activation of irrelevant information after a 1-s interval between sentence context and test picture (Gernsbacher & Faust, 1991; Gernsbacher et al., 1990; Merrill, Sperber, & McCauley, 1981).

However, to experience interference, one must have access to both meanings (Syssau et al., 2000) and must be aware of contextually relevant information (Merrill et al., 1981). Therefore, if children fail to demonstrate contextual facilitation, they would not be expected to experience interference (Merrill et al., 1981).

Method

Participants

The same children that took part in Experiment 1 completed the current experiment.

Materials

Several modifications were made to the Gernsbacher et al. (1990, Experiment 4) and Gernsbacher and Faust (1991, Experiment 4) materials to test both contextual facilitation and suppression of irrelevant information in the same participants. First, because children with LI frequently have concomitant reading difficulties (Bishop & Snowling, 2004), sentences were presented verbally and a picture verification task was used. Second, the interstimulus interval (ISI) was not varied. In Gernsbacher's work, varying the ISI did not affect contextual facilitation; both good and poor comprehenders benefited from context immediately as well as after a delay. In the

suppression conditions, differences between skill groups were seen only at the longer ISI. Furthermore, Merrill et al. (1981) suggested that a 1000-ms interval between sentence context and test word tapped the period of active comprehension in which context-specific selection of meaning was likely to occur rather than immediate lexical access. Therefore, the test picture appeared at 1000 ms from sentence offset in both conditions.

In this experiment, 22 sets of sentence stimuli were devised, using the same words and picture materials as were used in Experiment 1 (listed in the Appendix). For each ambiguous word, eight sentences were constructed, for a total of 176 sentences. Each sentence contained four or five simple words. Half of the sentences were used in the contextual facilitation condition, as shown in Table 3, and half were used in the suppression condition, as shown in Table 4.

In the facilitation condition, all sentences ended with ambiguous words. Each neutral sentence allowed either meaning of the ambiguous word. Each biased sentence manipulated the sentence context by altering the verb (Vu, Kellas, & Paul, 1998) so that only one meaning was acceptable (e.g., “He *ran* from the bank” vs. “He *fished/stole* from the bank”). Both neutral and biased sentences were followed by congruent pictures so that a correct response would be “yes.” If participants were sensitive to the semantic constraints of the verb, they should be more accurate and faster to press “yes” in the biased condition.

For the suppression condition, two sentence pairs were constructed for the dominant and subordinate meanings of the ambiguous words. In each pair, the verbs were identical and biased toward one meaning of the ambiguous word. The sentence pairs differed only in their final words; half of the sentences ended with an ambiguous word (e.g., “He stole from the *bank*”), and half ended with an unambiguous word (e.g., “He stole from the *shop*”). The pictures following these sentences were

Table 3
Sentences and pictures used in the contextual facilitation condition

Sentence type	Sentence	Picture
Neutral (dominant)	Bill ran from the BANK.	MONEY
Neutral (subordinate)	Bill ran from the BANK.	RIVER
Biased (dominant)	Bill stole from the BANK.	MONEY
Biased (subordinate)	Bill fished from the BANK.	RIVER

Note. The correct response to each picture is “yes.”

Table 4
Sentences and pictures used in the suppression condition

Sentence type	Sentence	Picture
Ambiguous (dominant)	Bill stole from the BANK.	RIVER
Unambiguous (dominant)	Bill stole from the SHOP.	RIVER
Ambiguous (subordinate)	Bill fished from the BANK.	MONEY
Unambiguous (subordinate)	Bill fished from the STREAM.	MONEY

Note. The correct response to each picture is “no.”

incongruent, so the correct response is “no.” In each case, although the picture was associated with the ambiguous word (e.g., *river*) it was not consistent with the meaning of the sentence. If children are sensitive to the sentence context and are able to suppress irrelevant meanings, they should reject the picture quickly regardless of the final word. However, if they have difficulty in suppressing irrelevant meanings, they should be slower to reject the picture following an ambiguous sentence.

All sentences were recorded by a male native speaker of British English. Sentences were divided into two lists with 88 sentences in each list. List A contained facilitation sentences for dominant meanings and suppression sentences in which the subordinate meanings were depicted. The reverse situation was presented in List B. Therefore, in each list, half of the items required a “yes” response. Sentences were pseudo-randomly ordered so that the same words/pictures did not follow one another. Presentation of the two lists was counterbalanced across participants, as was the order of items in each list. In most cases, the two lists were presented on different days. However, in the rare instances when testing had to be completed in one session, an interval of 30 min, during which other activities took place, separated each presentation.

Procedure

Once again, the experiment was run on a laptop computer using E-prime software. Children were instructed to listen to each sentence (presented through Sennheiser headphones) and then determine whether the picture that followed “fit” the meaning of the whole sentence. Children were asked to respond quickly but carefully, making sure that they had thought about the whole sentence. Eight practice items were given at the start of each list, and feedback on performance was given.

A “Ready?” screen initiated each trial and was under the experimenter’s control. This ensured that children were paying attention and allowed brief rests between trials where necessary. When the experimenter pressed the space bar, a fixation cross was displayed in the middle of the screen as the sentence was played through the headphones. At sentence offset, there was a 1000-ms delay and then a color picture appeared. The picture remained on the screen for 5 s or until the child pressed a button. Following the button press, the “Ready?” screen reappeared for the next trial. Both accuracy and RT were recorded.

Results

Analysis

All analyses were conducted using ANOVA unless otherwise stated. Analyses were then rerun using NVIQ as a covariate. This did not change the pattern of performance; therefore, only the ANOVA results are reported below. The effect size η^2 , which estimates the proportion of total variance accounted for by the independent variable, is reported where informative.

Before responses in each condition were considered, it was desirable to control for a possible bias toward “yes” responses and impulsive responses that could lead to chance performance. Because all facilitation items required a “yes” response and all suppression items required a “no” response, it was necessary to consider the pattern of response across the entire data set rather than in each condition. Initial exploration revealed that five participants (one from each clinical group and two from the TD group) failed to score on the subordinate suppression items (always responding “yes” to ambiguous items). It was believed that their responses to other items might be unreliable; therefore, they were excluded from further analysis. Signal detection theory provides a more stringent method for controlling for chance performance using the A' statistic. It compares the proportion of hits (correct “yes” and “no” responses) in relation to false alarms (pressing “yes” when a “no” response is required) and is calculated using the following formula (Gelfand, 1998):

$$A' = (0.5 * x * y) + [y * (1 - x)] + [0.5 * (1 - x) * (1 - y)],$$

where x equals the proportion of false alarms and y equals the proportion of hits. An A' score of 1.0 would indicate perfect performance, whereas a score of 0.5 would indicate a “yes” response to every item (chance performance).

Mean A' scores for each group are illustrated in Fig. 1. The figure demonstrates that scores for all groups were above chance. A one-way ANOVA on the A' scores revealed a modest but significant group difference, $F(3,87) = 8.44$, $p < .001$, $\eta^2 = .23$. Post hoc comparisons revealed that the LI and ALI groups did not differ from one another and had significantly lower scores than did the ASO and TD groups, which also did not differ from one another, $ps < .02$. This indicates that although the two groups with language impairment (LI and ALI) were performing above chance, they were significantly less accurate overall than were the two groups without language impairment (ASO and TD). The facilitation and suppression conditions were analyzed separately to reveal the source of this inaccuracy.

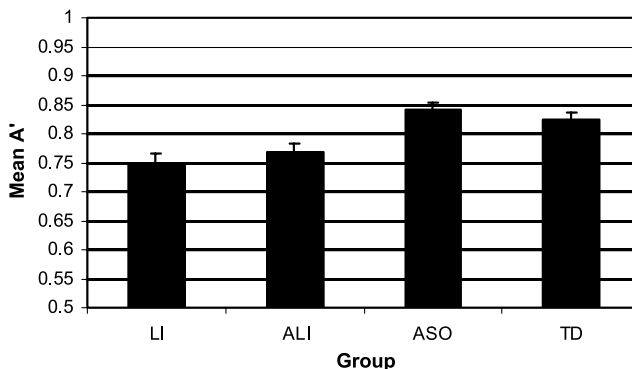


Fig. 1. Mean A' scores per group. A score of 1.0 indicates perfect performance, whereas a score of 0.50 equals chance performance. Error bars are standard errors.

Contextual facilitation

Table 5 shows the mean numbers of correct responses and mean RTs for ambiguous items in neutral and biased sentence contexts. Two 4 (Group) \times 2 (Context Type: biased vs. neutral) \times 2 (Word Type: dominant vs. subordinate) repeated measures ANOVAs were conducted, one using number of correct items as the dependent variable and one using mean RTs as the dependent variable.

Accuracy

There were main effects of context type, $F(1, 87) = 363.90$, $p < .001$, $\eta^2 = .81$, $F(1, 21) = 23.25$, $p = .002$, $\eta^2 = .53$, and word type, $F(1, 87) = 244.90$, $p < .001$, $\eta^2 = .74$, $F(1, 21) = 5.57$, $p = .028$, $\eta^2 = .21$, indicating that participants were more accurate in the biased condition than in the neutral condition and that they were more accurate with dominant meanings than with subordinate meanings. The Context Type \times Word Type interaction was significant, $F(1, 87) = 115.15$, $p < .001$, $\eta^2 = .56$, $F(1, 21) = 5.82$, $p = .03$, $\eta^2 = .22$, indicating that the biasing context exerted a greater effect on subordinate items than on dominant meanings. The main effect of group was significant by items only, $F(3, 87) = 1.43$, $p > .10$, $F(3, 63) = 5.41$, $p = .002$, $\eta^2 = .21$. However, the Group \times Context Type interaction was significant by both subject and items analyses, $F(3, 87) = 6.23$, $p < .001$, $\eta^2 = .18$, $F(3, 63) = 11.18$, $p < .001$, $\eta^2 = .35$. No other interaction terms were significant.

The significant Group \times Context Type interaction indicates that the amount of contextual facilitation differed across the four groups. This was investigated by calculating a contextual facilitation score for each child. This was derived by subtracting the total correct score in the neutral conditions (dominant + subordinate) from the total correct score in the biased condition. Therefore, children who benefited more from the biasing contexts would have greater difference scores than would chil-

Table 5

Mean numbers correct (maximum 22), RTs (in milliseconds), and facilitation scores per group

		<i>n</i>	Dominant		Subordinate		Total facilitation
			Bias	Neutral	Bias	Neutral	
Accuracy	LI	19	18.79 (2.46)	18.63 (2.31)	17.16 (3.02)	8.89 (3.26)	8.42 (6.47) ^a
	ALI	27	19.07 (2.09)	16.89 (3.33)	17.19 (3.52)	9.81 (4.25)	9.56 (7.43) ^{ac}
	ASO	19	19.47 (1.68)	16.00 (3.11)	19.42 (1.77)	8.84 (3.24)	14.05 (4.06) ^{bc}
	TD	26	20.12 (1.42)	17.46 (2.44)	19.88 (1.37)	8.23 (3.49)	14.31 (3.76) ^b
RTs	LI	19	1057 (378)	1103 (372)	1206 (319)	1410 (392)	251 (259)
	ALI	27	1173 (389)	1262 (390)	1315 (376)	1612 (476)	386 (369)
	ASO	19	1189 (434)	1374 (517)	1258 (402)	1548 (533)	476 (346)
	TD	26	1045 (330)	1185 (380)	1108 (447)	1382 (480)	414 (291)

Note. Standard deviations are in parentheses. Total Facilitation = (Dominant Bias + Subordinate Bias) – (Dominant Neutral + Subordinate Neutral). Larger facilitation scores signify greater contextual facilitation. Items with different superscripts in the same column are significantly different at $p < .05$. LI, language impairment; ALI, autistic spectrum disorder plus language impairment; ASO, autistic spectrum disorder and verbal abilities within the normal range; TD, typically developing; and RT, response time.

dren who benefited less from context. As illustrated in Fig. 2, this revealed a significant effect of group, $F(3, 87) = 6.23$, $p < .001$, $F(3, 63) = 11.18$, $p < .001$. Planned comparisons revealed that both the LI and ALI groups experienced less contextual facilitation than did the TD groups: LI versus TD, $p < .01$; ALI versus TD, $p = .02$. Similarly, the ASO group showed significantly more facilitation than did the LI group, $p = .02$, and showed a trend for more facilitation than the ALI group, $p = .06$. The ASO and TD groups did not differ from one another.

Chapman, Chapman, Curran, and Miller (1994) urged caution in interpreting such difference scores. They argued that individual and group differences in such scores are heavily influenced by initial accuracy, leading to the somewhat paradoxical finding that children and clinical populations may show heightened priming effects. In other words, children who are less accurate to begin with may appear to benefit more from context simply because they have more room for improvement. This did not appear to be the case here given that the TD and ASO groups demonstrated the greatest amount of facilitation despite having better overall accuracy scores. One possibility is that the children with LI (the LI and ALI groups) did not show as much contextual facilitation because they simply were not as familiar with both meanings of the ambiguous words. To control for this possibility, errors made in Experiment 1 were examined. If a child did not know both the dominant and subordinate meanings of a word, sentences containing that word were eliminated from the current analysis. Therefore, this analysis uses proportion of items correct, rather than raw score, as the dependent variable.

A 4 (Group) \times 2 (Context Type: biased vs. neutral) \times 2 (Word Type: dominant vs. subordinate) repeated measures ANOVA revealed a broadly similar pattern of results, with a main effect of context type, $F(1, 87) = 153.67$, $p < .001$, $F(1, 21) = 25.77$, $p < .001$. The main effect of word type was significant by subjects only, $F(1, 87) = 29.37$, $p < .001$, $F(1, 21) = 2.89$, $p > .10$, with more correct responses to dominant meanings. Although the main effect of group was not significant by either subjects or items, $F(3, 87) = 3.54$, $p = .07$, the Group \times Context Type interaction remained reliable, $F(3, 87) = 3.95$, $p = .01$, $F(3, 63) = 7.07$, $p = .001$. In this more conservative analysis, the LI group continued to demonstrate less contex-

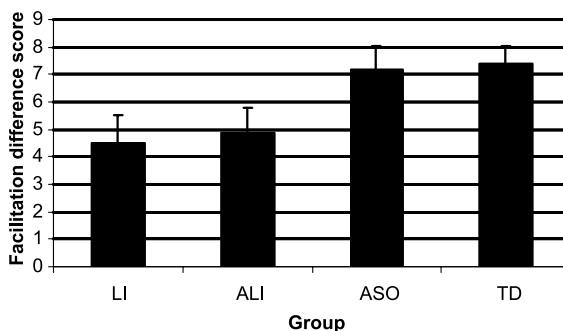


Fig. 2. Mean contextual facilitation scores per group. Error bars are standard errors.

tual facilitation than did either the ASO group or the TD group, $ps < .05$. There were no other group differences. None of the other interaction terms was significant.

These results suggest that although children with language impairment (the LI and ALI groups) were sensitive to the contextual constraints (as evidenced by their higher accuracy scores in biased vs. neutral conditions), they did not process context as efficiently as did the two groups without LI. Even when knowledge of the ambiguous word meanings was controlled, the LI group continued to experience more difficulties with contextual facilitation than did the ASO and TD groups.

Response times

With respect to RTs, Fig. 3 illustrates the spread and variability of scores for each group across conditions. Although the range was large, the variance across groups was roughly equal. Following Ulrich and Miller (1994), extreme RTs were not eliminated because truncating the data in this way may introduce bias. Therefore, RTs were analyzed for all correct responses. Analyses of both mean and median RTs (which are likely to be more stable) were conducted, but because these analyses produced the same pattern of results, only mean RTs are reported.

A 4 (Group) \times 2 (Context Type: biased vs. neutral) \times 2 (Word Type: dominant vs. subordinate) repeated measures ANOVA revealed significant main effects of context type, $F(1, 87) = 124.18$, $p < .001$, $\eta^2 = .59$, $F(1, 21) = 46.57$, $p < .001$, $\eta^2 = .69$, and Word Type, $F(1, 87) = 19.85$, $p < .001$, $\eta^2 = .19$, $F(1, 21) = 10.03$, $p < .001$, $\eta^2 = .32$, such that all groups were faster to respond in the biased condition and to dominant items (Table 5). The main effect of group was significant by items only, $F(1, 87) = 1.46$, $p > .10$, $F(3, 63) = 20.68$, $p < .001$, $\eta^2 = .50$. Here the LI and TD groups were faster to respond than were the ALI and ASO groups¹. None of the interaction terms was significant by either subjects or items: Group \times Word Type, $F(1, 87) < 1$, $F(3, 63) = 1.39$, $p > .10$; Group \times Context Type, $F(1, 87) = 1.68$, $p > .10$, $F(3, 63) = 2.39$, $p = .09$, $p > .10$; Group \times Context Type \times Word Type, $F(1, 87) < 1$, $F(3, 63) < 1$. When analysis was restricted to only those items in which both the dominant and subordinate meanings were known (Experiment 1), a similar pattern emerged with main effects of type and context: Word Type, $F(1, 87) = 29.37$, $p < .001$, $F(1, 21) = 9.08$, $p = .007$; Context Type, $F(1, 87) = 153.67$, $F(1, 21) = 50.37$, $ps < .001$, such that all groups responded more quickly to dominant meanings and biased sentences. The main effect of group was significant by items only, $F(1, 87) < 1$, $F(3, 63) = 13.62$, $p < .001$. Again, the LI and TD groups were quicker to respond than were the ALI and ASO groups. None of the interaction terms was significant.

Summary

All groups were more accurate and faster when judging dominant meanings as opposed to subordinate meanings. In addition, all groups showed evidence of contextual facilitation in terms of increased accuracy and faster RTs to pictures following biased, as opposed to neutral, sentence contexts. However, the groups with language

¹ Analysis of RTs to errors revealed no significant group differences, suggesting that there was not a speed–accuracy trade-off for the LI group.

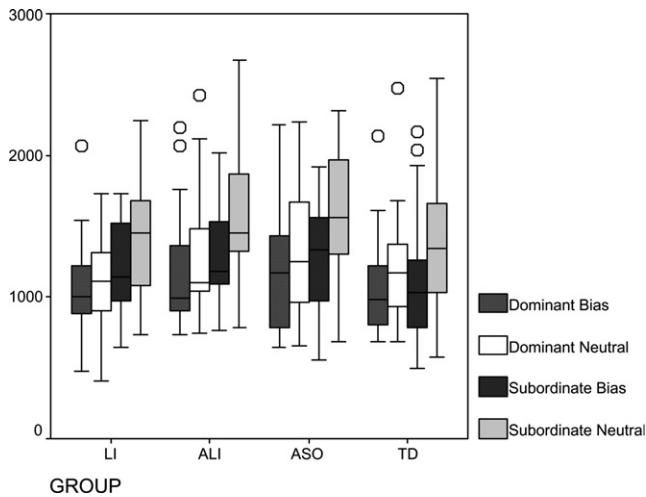


Fig. 3. Box plots depicting response times (in milliseconds) on contextual facilitation items in neutral and biased conditions. The boxes represent the middle 50% of scores, with the black line in the middle of the box illustrating the median. The whiskers represent the highest and lowest values. Circles are extreme values.

impairment (the LI and ALI groups) did not show the same degree of contextual facilitation in terms of accuracy. These findings are at odds with weak central coherence accounts of ASD that would predict deficits in contextual facilitation for all individuals with ASD across the spectrum of ability. The results of this study clearly show that children with autistic features who have good verbal skills (as measured by standardized tests of language ability) experience the same amount of contextual facilitation as do their TD peers. Children in the ALI group, who had impaired language abilities, were less efficient at processing context. When knowledge of multiple meanings was controlled, facilitation scores improved. However, it is important to realize that the contextual constraints provided by the verbs did not help children with language impairment to deduce the possible meanings of the sentence final ambiguous words.

Inefficiencies in contextual processing were not limited to children with autistic features. Nonautistic children with LI performed similarly to the ALI group in terms of accuracy. Furthermore, even when only words for which they knew both meanings were included in the analysis, the contextual facilitation scores for the LI group continued to lag behind not only their TD peers but also children with autistic features without LI (the ASO group). This suggests further difficulty in understanding verb meanings and linking ideas together in a short sentence context.

Suppression

Gernsbacher and colleagues (e.g., Gernsbacher & Faust, 1991) demonstrated that adult poor comprehenders showed adequate contextual facilitation but were poor at

suppressing irrelevant information. However, in the current study, children with poor language comprehension, regardless of autistic status (the LI and ALI groups), were not as efficient in using contextual constraints as were their peers without language impairment (the ASO and TD groups). Adequate contextual facilitation is necessary because one cannot determine what is contextually irrelevant without having a sense of what is contextually relevant (Merrill et al., 1981; Syssau et al., 2000). Therefore, contrary to Gernsbacher and colleagues, a strong prediction can be made that children with language impairment will *not* show evidence of interference.

In line with Gernsbacher and colleagues, children in the TD group were not expected to show interference given the 1000-ms delay between sentence offset and picture. The performance of the ASO group was less predictable. However, because children in the ASO group score within normal limits on tests of sentence processing and experience the same degree of contextual facilitation as do their TD peers, they were expected to perform similarly to the TD group.

Accuracy

Suppression cannot occur without activation (Syssau et al., 2000); therefore, only those words for which children knew both meanings in Experiment 1 were included in this analysis. Table 6 shows the mean proportions correct and mean RTs for correct responses to both ambiguous and unambiguous sentences. On unambiguous items, it is apparent that all groups are performing near ceiling and have little difficulty in rejecting erroneous items. In contrast, performance is poorer and more variable in the ambiguous condition, meaning that children incorrectly accept pictures depicting the alternative meanings of the ambiguous words (e.g., accept *money* when the sentence context biases *river*). A 4 (Group) \times 2 (Ambiguity: ambiguous vs. unambiguous) \times 2 (Word Type: dominant vs. subordinate) repeated measures ANOVA

Table 6

Mean (SD) proportion correct, RTs (in milliseconds) and interference scores, both dominant and subordinate conditions

		<i>n</i>	Dominant		Subordinate		Total interference
			A	U	A	U	
Accuracy	LI	19	.52 (.22)	.87 (.15)	.60 (.18)	.96 (.06)	.71 (.32)
	ALI	27	.60 (.24)	.91 (.08)	.63 (.26)	.96 (.08)	.64 (.42)
	ASO	19	.78 (.12)	.97 (.05)	.78 (.14)	.98 (.03)	.39 (.17)
	TD	26	.73 (.18)	.97 (.09)	.75 (.14)	.98 (.04)	.45 (.26)
RTs	LI	19	1369 (558)	1234 (382)	1501 (538)	1311 (354)	.14 (.19)
	ALI	27	1667 (621)	1381 (439)	1475 (394)	1431 (457)	.13 (.21)
	ASO	19	1393 (511)	1308 (453)	1455 (523)	1335 (438)	.08 (.12)
	TD	26	1340 (414)	1149 (325)	1368 (462)	1246 (401)	.13 (.13)

Note. Standard deviations are in parentheses. Total Interference RTs = [(Dominant Ambiguous + Subordinate Ambiguous) – (Dominant Unambiguous + Subordinate Unambiguous)] / (Dominant Unambiguous + Subordinate Unambiguous). A, ambiguous; U, unambiguous; LI, language impairment; ALI, autistic spectrum disorder *plus* language impairment; ASO, autistic spectrum disorder and verbal abilities within the normal range; TD, typically developing; and RT, response time.

confirmed a significant effect of ambiguity, $F(1,87) = 264.83$, $p < .001$, $\eta^2 = .75$, $F(1,21) = 95.23$, $p < .001$, $\eta^2 = .82$, with greater accuracy on unambiguous items. There was also a main effect of group, $F(3,87) = 8.09$, $p < .001$, $\eta^2 = .22$, $F(3,63) = 44.39$, $p < .001$, $\eta^2 = .68$. Post hoc analyses revealed that the TD and ASO groups had more correct responses than did the LI and ALI groups, $ps < .05$. There was a small but significant effect of word type by subjects only, $F(1,87) = 9.89$, $p = .002$, $\eta^2 = .10$, $F(2,1) < 1$, with slightly more correct responses to subordinate items. The Group \times Ambiguity interaction was significant, $F(3,87) = 4.92$, $p = .003$, $\eta^2 = .15$, $F(3,63) = 22.85$, $p < .001$, $\eta^2 = .52$. Although there were minimal group differences on unambiguous items, $F(3,87) = 4.16$, $p < .01$, with the LI group having significantly lower scores than the ASO group, $p = .02$, the significant differences on ambiguous items were more striking, $F(3,87) = 7.20$, $p < .001$, with both the LI and ALI groups making more errors than the TD and ASO groups: LI versus TD and ASO, $ps < .007$; ALI versus TD, $p = .07$; ALI versus ASO, $p = .01$. None of the other interaction terms was significant.

Response times

Again, there was large within-group variation but equal variance across groups. Data were analyzed using untrimmed means on correct items only (Ulrich & Miller, 1994). Both means and medians were analyzed, but both produced the same pattern of results, so only mean RTs are reported. Again, only those items where both dominant and subordinate meanings were known were included.

There was a main effect of ambiguity, $F(1,86) = 36.28$, $p < .001$, $\eta^2 = .29$, $F(1,20) = 31.44$, $p < .001$, $\eta^2 = .61$, such that all groups were significantly faster to reject pictures following unambiguous sentences than following ambiguous sentences. The main effect of word type was not significant, $F(1) < 1$ and $F(2) < 1$. The effect of group was significant by items only, $F(3,86) = 1.27$, $p > .10$, $F(3,60) = 9.28$, $p < .001$, $\eta^2 = .12$, with the ALI group responding more slowly than the ASO group, $p = .01$, and the TD group, $p = .001$. The Group \times Ambiguity and Group \times Word Type interaction terms were also significant by items but not by subjects: Group \times Ambiguity, $F(3,60) = 3.37$, $p = .02$; Group \times Word Type, $F(3,60) = 3.39$, $p = .02$. These interaction terms were not analyzed further.

An interference measure was computed by subtracting RTs (dominant and subordinate meanings combined) to reject pictures following unambiguous words (e.g., *shop*) from RTs to reject pictures following ambiguous items (e.g., *bank*) and dividing by speed on unambiguous items (this corrects for general speed). Large positive numbers are, therefore, indicative of more interference from the irrelevant meanings of the ambiguous words in the ambiguous conditions. All groups had positive mean scores, and there were no significant group differences on this measure, with both F s < 1 .

Summary

All groups were quick to correctly reject irrelevant picture meanings that appeared after unambiguous sentences (e.g., “Bill stole from the *shop*” followed

by a picture of a *river*). However, performance was far more variable, and RTs were slower, when pictures followed ambiguous sentences (e.g., “Bill stole from the *bank*” followed by a picture of a *river*). In particular, groups of children with structural language impairment (the LI and ALI groups) made significantly more errors than did their peers without language impairment. These results could be taken as evidence of an inefficient suppression mechanism characterizing children with LI. However, performance on the contextual facilitation items suggests a different story. Here the LI and ALI groups experienced less contextual facilitation; that is, they were less able to link ideas together in sentences to disambiguate meanings. Therefore, it was predicted that they would not show evidence of interference. Interference-like errors could be made if children failed to take prior context into account and matched only the final words of the sentences to the pictures. Children would then be more likely to respond “yes” to ambiguous items, resulting in a greater number of errors. In this way, children with language impairment appeared to perform the task as if it were a word–picture verification task rather than a sentence–picture verification task. This would result in ceiling performance on unambiguous items (e.g., *river–money* is clearly incorrect) but in closer to chance performance on ambiguous items (e.g., *bank–money* would be correct in a word–picture matching task). This result is in keeping with the findings of Merrill et al. (1981). Using a modified Stroop task, they found that children with poor reading comprehension responded to test words following sentence contexts in a similar fashion to how they responded to test words following words in isolation. This suggests that the poor comprehenders failed to take the sentence context into account and were more reliant on the “objective salience” (or dominant meanings) of the individual lexical items.

TD children also seemed to experience some degree of interference, with more accurate and faster responses to unambiguous items. Because they did show knowledge of both word meanings and adequate contextual facilitation, it is more plausible that both meanings were activated in the suppression condition. These findings are at odds with those of Gernsbacher and colleagues, who found that despite initial activation of both word meanings, after a delay of 850 ms the skilled adult comprehenders no longer experienced interference of contextually irrelevant information. This suggests the possibility that the mechanism of suppression may develop with age (Lorsbach, Katz, & Cupak, 1998). Other investigators have found that school-age children are able to suppress irrelevant information after a delay of 1 s (Merrill et al., 1981). However, these studies have focused on poor reading comprehension and have used different techniques such as a Stroop task. Therefore differences between studies may be a reflection of the different tasks used to investigate suppression.

General discussion

The current experiments investigated three sources of comprehension breakdown in two “distinct” clinical populations, SLI and ASD, to determine whether the

comprehension deficits reported for these children occur for different reasons. A novel approach to group selection was adopted to tease apart aspects of comprehension that are characteristic of autistic performance at any level of verbal ability and those that are part and parcel of an impairment in fundamental aspects of language—semantic knowledge and sentence processing.

The first possibility outlined in the introduction was that children with ASD, regardless of language status, would be impaired in their ability to extract meaning from context. This was not the case. Only the children with ASD who had concomitant structural language impairment showed deficient contextual facilitation. Higher functioning children with ASD who had structural language scores within the normal range did not differ from their TD peers in the use of context to resolve lexical ambiguities. Furthermore, deficits in contextual processing were not limited to children with ASD. Children with LI who did not present with current autistic features performed similarly to the ALI group in deriving less facilitation from context.

This finding is not unique. [Snowling and Frith \(1986\)](#) found, in their autistic sample, that the ability to produce the correct pronunciation of homographs in context was related to verbal ability. Those children with good verbal ability performed as well as did controls. Furthermore, low verbal ability impeded the ability to use context in both autistic children and nonautistic children with intellectual impairment. These findings further highlight the importance of controlling for structural language skills when evaluating cognitive deficits in autism (cf. [Bishop & Norbury, 2005](#)).

The second possibility outlined in the introduction was that the two groups of children with core language deficits, one with additional autistic features and one without, would show deficits in processing context in order to resolve lexical ambiguities. This was indeed the case. Experiment 1 illustrated that the LI and ALI groups were less familiar with the subordinate meanings of ambiguous words, reflecting their generally poorer vocabulary levels. Experiment 2 demonstrated that these groups did not experience as much contextual facilitation as did their more linguistically able peers, even when knowledge of both meanings was accounted for. What contributes to their difficulties with contextual processing?

One possibility is the impoverished semantic knowledge of the children with language impairment. The meanings of the ambiguous test words were constrained by manipulating the verbs. Investigations of skilled adult language users have shown that individuals rapidly activate semantic information about the verbs they encounter, including thematic role information such as typical agents and patients ([Ferrelli, McRae, & Hatherell, 2001](#)). There is a dearth of experimental investigation into the understanding of thematic roles by children with SLI or ASD. However, there has been some suggestion that children with SLI have difficulty in acquiring new verbs ([Windfuhr, Faragher, & Conti-Ramsden, 2002](#)) and have a tendency to use less specific verbs in spontaneous language ([Rice & Bode, 1993](#)), suggesting that they might not be sensitive to the semantic constraints imposed by the verbs. Future investigations should explore verb knowledge in both SLI and ASD as well

as in other sources of sentence constraint such as the noun phrase (e.g., “The thief ran from the bank”) (Vu et al., 1998). Clearly, children with language difficulties, with or without ASD, need explicit instruction in the multiple meanings of ambiguous words and how to link ideas in sentences/discourse together to interpret meaning correctly.

A second source of difficulty may be children’s memory for prior linguistic context. Poor sentence memory would result in important contextual information being unavailable at the point of disambiguation. If this were the case, children might treat the task as a word-to-picture matching task, as suggested by the LI and ALI groups’ responses to suppression items. There is ample evidence that children with SLI have reduced memory for sentences, particularly when the processing demands of the task are high (Conti-Ramsden, Botting, & Faragher, 2001; Montgomery, 2000a, 2000b). Although children with autism are frequently reported to have good rote memory skills, there is some suggestion that their memory for structured information, such as sentences, may be less impressive (Minschew & Goldstein, 2001).

The LI group had significantly lower NVIQ scores than did the ASO and TD groups, raising the possibility that the LI group’s poorer performance may be due in part to differences in nonverbal ability. This did not appear to be the case. Partialling out NVIQ scores in a statistical analysis did not alter the pattern of results. Furthermore, the ALI group did not differ significantly from any of the other groups with respect to nonverbal ability. The fact that the ALI group’s performance more closely resembled that of the children with LI suggests that it is differences in core language ability that are more typically associated with contextual processing. Further investigation into the role of nonverbal ability in *suppressing* irrelevant information is warranted.

A final source of difficulty considered here was the ability to suppress contextually irrelevant information. With respect to the suppression condition, children in the LI and ALI groups made errors that looked like they were failing to suppress irrelevant meanings. However, given that they were less able to appreciate what was contextually relevant, these errors suggest that the children were processing the meanings of the individual words rather than linking words together. This does not mean that children with language impairment have efficient suppression mechanisms. Errors made by children with SLI on inferencing tasks do suggest that irrelevant meanings may be activated over a long period of time (Norbury & Bishop, 2002). However, it is difficult to determine at what point these errors arise—as the sentences unfold or at the point when children are asked questions about what they have heard.

Similar questions arise in interpreting the errors of the TD and ASO groups on suppression items. They were slower to respond to ambiguous versus unambiguous sentences, and Gernsbacher has suggested that this is evidence of interference. However, Long, Seely, and Oppy (1999), in an investigation of children’s reading comprehension processes, argued that these effects are inherent in the task demands. In their study, poor comprehenders showed evidence of suppression difficulties when performing a meaning judgment task (similar to the one used in the current study) but performed as well as skilled comprehenders on a lexical decision

task that arguably had fewer metacognitive demands. They argued that the mechanism of suppression might not be automatic; instead, it might be under conscious control and subject to available processing resources such as attention and memory.

An important avenue for future research will be investigating individual differences in language skill (particularly verb knowledge) and contextual processing using online methodologies such as eye tracking. Such methods reduce the metacognitive demands and memory load of the tasks used. They also provide an index of comprehension as it happens, further defining the timing and source of comprehension deficit in children with communication disorders (Nation, Marshall, & Altmann, 2003).

Conclusions

The two experiments reported in this study provide further evidence that children with language impairment have difficulty in integrating ideas together in context to resolve ambiguous messages. These difficulties may be due in part to core deficits in semantic knowledge and sentence processing. However, they may also reflect deficits in memory and attention skills that are necessary for developing mental representations of spoken input (Bishop, 1997).
















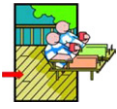






Contrary to the predictions of weak central coherence theory, deficits in contextual processing were not seen across the autistic spectrum. Those individuals who had core language scores within the normal range performed similarly to typically developing peers in both contextual facilitation and suppression of irrelevant information. These experiments highlight the importance of distinguishing between language ability and autistic symptoms in the ASD population. Only those with structural language impairment show evidence of weak central coherence in the verbal domain.

Acknowledgments

This work was completed while the author was in possession of a Wellcome Trust Prize Studentship and was submitted in partial fulfillment of a Doctor of Philosophy degree at the University of Oxford. I thank Dorothy Bishop and Kate Nation for their insightful comments and helpful suggestions on an earlier draft of this article. A very warm thank you is extended to all of the children and families who took part in this study and to the following schools that assisted in recruitment of participants and hosting this research: Barley Hill Primary, Blossom House, Chinnor Resource Base, Donnington Language Resource Base, Gap House, Hope Lodge, John Horniman School, Link Primary and Secondary schools, Moor House, More House, Oxted County School, St. Catherine's School, St. Dominic's School, Sir Thomas Abney Language Resource Base, and Whitefields School.












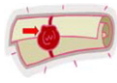










Appendix

Ambiguous words and pictures with dominance ratings (percentages of total responses)

Word	Dominant meaning	Rating	Subordinate meaning	Rating
Bank		92.31		5.13
Bat		51.28		46.15
Bow		58.97		25.64
Bulb		74.36		23.08
Coach		82.05		15.38
Cold		89.74		10.26
Cup		94.87		5.13
Deck		46.15		7.69
Glass		53.85		35.90
Horn		61.54		28.21
Jam		89.74		7.69

(continued on next page)

Appendix (continued)

Word	Dominant meaning	Rating	Subordinate meaning	Rating
Match		66.67		30.77
Mouse		87.18		10.26
Nut		84.62		5.13
Organ		76.92		20.51
Palm		56.41		43.59
Seal		74.36		15.38
Spade		92.31		2.56
Speaker		69.23		25.64
Straw		61.54		30.77
Trunk		51.28		10.26
Wave		66.67		33.33

Sentences used in Experiment 2. Manipulations of the verb (neutral/dominant/subordinate) and ambiguous/unambiguous endings are also shown.

John ran/stole/fished from the *bank/shop/river*.
 Peter found/hurt/swung the *bat/bird/stick*.
 Julie bought/strung/wore the *bow/guitar/tie*.
 Peter bought/changed/planted the *bulb/light/seed*.
 Bill saw/drove/listened to the *coach/bus/teacher*.
 Jill hated/felt/treated the *cold/weather/flu*.
 Dan carried/broke/won the *cup/pot/trophy*.
 Tom looked at/shuffled/swept the *deck/cards/patio*.
 Jill broke/wiped/filled the *glass/window/tumbler*.
 Tom grabbed/sounded/covered the *horn/trumpet/nose*.
 Peter hated/ate/left the *jam/toast/traffic*.
 Sally watched/lit/played the *match/fire/game*.
 Steve bought/chased/connected the *mouse/rat/computer*.
 Paul lost/ate/twisted the *nut/almond/screw*.
 Ann recognized/tuned/transplanted the *organ/piano/heart*.
 Lisa touched/squeezed/climbed the *palm/lemon/tree*.
 Jill liked/trained/broke the *seal/dolphin/code*.
 Paul picked up/dug with/played the *spade/shovel/ace*.
 Julie listened to/connected/thanked the *speaker/stereo/teacher*.
 Peter chewed/bailed/sucked the *straw/hay/drink*.
 Peter touched/stroked/opened the *trunk/face/case*.
 Julie watched/surfed/returned the *wave/ocean/kiss*.

References

- Bartak, L., Rutter, M., & Cox, A. (1975). A comparative study of infantile autism and specific developmental receptive language disorders: I. The children. *Journal of Autism and Childhood Schizophrenia*, 7, 383–396.
- Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry*, 175, 444–451.
- Bishop, D. V. M. (1979). Comprehension in developmental language disorders. *Developmental Medicine and Child Neurology*, 21, 225–238.
- Bishop, D. V. M. (1997). *Uncommon understanding: Development and disorders of language comprehension in children*. Hove, UK: Psychology Press.
- 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., & Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: A study using standardised diagnostic instruments. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43, 917–931.
- Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments in relation to autistic symptomatology 2: Response inhibition. *Autism*, 9(1), 29–43.
- Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*, 130(6), 858–886.

- Chapman, L. J., Chapman, J. P., Curran, T. E., & Miller, M. B. (1994). Do children and the elderly show heightened semantic priming? How to answer the question. *Developmental Review*, 14, 159–185.
- Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42, 741–748.
- Dunn, L. M., Dunn, L. M., & Whetton, C. W. (1997). *The British picture vocabulary scale* (2nd ed.). Windsor, UK: NFER–Nelson.
- Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language*, 44, 516–547.
- Frith, U. (1989). *Autism: Explaining the enigma*. Oxford, UK: Basil Blackwell.
- Frith, U., & Happé, F. (1994). Autism: Beyond theory of mind. *Cognition*, 50, 115–132.
- Gelfand, S. A. (1998). *Hearing: An introduction to psychological and physiological acoustics* (3rd ed.). New York: Marcel Dekker.
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum.
- Gernsbacher, M. A., & Faust, M. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 245–262.
- Gernsbacher, M. A., & Robertson, R. (1999). The role of suppression in figurative language comprehension. *Journal of Pragmatics*, 31, 1619–1630.
- Gernsbacher, M. A., Varner, K. R., & Faust, M. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 430–445.
- Happé, F. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, 15, 1–12.
- Happé, F. (1999). Autism: Cognitive deficit or cognitive style?. *Trends in Cognitive Sciences*, 3, 216–222.
- Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, 24, 189–223.
- Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger's syndrome: Is local coherence impaired?. *Cognition*, 71, 149–185.
- Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger's syndrome: Is global coherence impaired?. *Psychological Medicine*, 30, 1169–1187.
- Kerbel, D., & Grunwell, P. (1998). A study of idiom comprehension in children with semantic–pragmatic difficulties: II. Between-groups results and discussion. *International Journal of Language and Communication Disorders*, 33, 23–44.
- Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, 16, 287–308.
- Long, D. L., Seely, M. R., & Oppy, B. J. (1999). The strategic nature of less skilled readers' suppression problems. *Discourse Processes*, 27, 281–302.
- Lopez, B., & Leekham, S. R. (2003). Do children with autism fail to process information in context?. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44, 285–300.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–685.
- Lorsbach, T. C., Katz, G. A., & Cupak, A. J. (1998). Developmental differences in the ability to inhibit the initial misinterpretation of garden path passages. *Journal of Experimental Child Psychology*, 71, 275–296.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676–703.
- McGregor, K. K., Newman, R. M., Reilly, R., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 45, 998–1014.
- Merrill, E. C., Sperber, R. D., & McCauley, C. (1981). Differences in semantic encoding as a function of reading comprehension skill. *Memory & Cognition*, 9, 618–624.

- Minschew, N. J., & Goldstein, G. (2001). The pattern of intact and impaired memory functions in autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42, 1095–1101.
- Montgomery, J. W. (2000a). Relation of working memory to off-line and real-time sentence processing in children with specific language impairment. *Applied Psycholinguistics*, 21, 117–148.
- Montgomery, J. W. (2000b). Verbal working memory and sentence comprehension in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 43, 293–308.
- Nation, K., Marshall, C. M., & Altmann, G. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology*, 86, 314–329.
- 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. (in press). The relation between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology*.
- Norbury, C. F., & Bishop, D. V. M. (2002). Inferential processing and story recall in children with communication problems: A comparison of specific language impairment, pragmatic language impairment, and autism. *International Journal of Language and Communication Disorders*, 37, 227–251.
- Ozonoff, S., & Miller, J. (1996). An exploration of right-hemisphere contributions to the pragmatic impairments of autism. *Brain and Language*, 52, 411–434.
- Rice, M. L., & Bode, J. V. (1993). GAPS in the verb lexicons of children with specific language impairment. *First Language*, 13, 113–131.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools.
- Semel, E. M., Wiig, E. H., & Secord, W. (2000). *Clinical evaluation of language fundamentals* (3rd U.K. ed.). London, UK: Psychological Corporation.
- Snowling, M. J., & Frith, U. (1986). Comprehension in “hyperlexic” readers. *Journal of Experimental Child Psychology*, 42, 392–415.
- Swaab, T., Brown, C., & Hagoort, P. (2003). Understanding words in sentence contexts: The time course of ambiguity resolution. *Brain and Language*, 86, 326–343.
- Syssau, A., Brouillet, D., & Groen, S. (2000). The study of inhibitory processes in aging with the Faust et al. paradigm (1997): Methodological issues. *Brain and Language*, 72, 150–157.
- Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language*, 32, 359–372.
- Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London*, 358, 303–314.
- Tyler, L. K., & Marslen-Wilson, W. (1981). Children's processing of spoken language. *Journal of Verbal Learning and Verbal Behavior*, 20, 400–416.
- Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General*, 123, 34–80.
- Vu, H., Kellas, G., & Paul, S. T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & Cognition*, 26, 979–1001.
- Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: Psychological Corporation.
- Windfuhr, K. E., Faragher, B., & Conti-Ramsden, G. (2002). Lexical learning skills in young children with specific language impairment. *International Journal of Language and Communication Disorders*, 37, 415–432.