Receptive Vocabulary and Semantic Knowledge in Children with SLI and Children with Down Syndrome

Glynis Laws\textsuperscript{a}, Josie Briscoe\textsuperscript{a}, Su-Yin Ang\textsuperscript{a}, Heather Brown\textsuperscript{a}, Ehab Hermena\textsuperscript{a} and Anna Kapikian\textsuperscript{a}

\textsuperscript{a}University of Bristol

Author Note

Glynis Laws, Josie Briscoe, Su-Yin Ang, Heather Brown, Ehab Hermena and Anna Kapikian, School of Experimental Psychology, University of Bristol.

Su-Yin Ang is now at Nanyang Technological University. Ehab Hermena is now at the University of Southampton. Anna Kapikian is now at the Institute of Education, London.

This research was completed with the support of the Wellcome Trust (project grant number 076520). Laura Bence, Alison Fisher, Stephanie Guillaume, Elizabeth Main, Emily Mason-Apps, and Catherine White contributed to assessments or data coding and entry. Some data used in the study were gathered for MSc dissertations completed by Heather Brown and Anna Kapikian.

Correspondence concerning this article should be addressed to Glynis Laws, School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol, BS8 1TU, UK. E-mail: glynis.laws@bristol.ac.uk
Abstract

Receptive vocabulary and associated semantic knowledge were compared within and between groups of children with Specific Language Impairment (SLI), children with Down syndrome (DS) and typically developing children. To overcome the potential confounding effects of speech or language difficulties on verbal tests of semantic knowledge, a novel task was devised based on picture-based semantic association tests used to assess adult patients with semantic dementia. Receptive vocabulary, measured by word-picture matching, of children with SLI was weak relative to chronological age and to nonverbal mental age but their semantic knowledge, probed across the same lexical items, did not differ significantly from that of vocabulary-matched typically developing children. By contrast, although receptive vocabulary of children with DS was a relative strength compared to nonverbal cognitive abilities ($p < .0001$), DS was associated with a significant deficit in semantic knowledge ($p < .0001$) indicative of dissociation between word-picture matching vocabulary and depth of semantic knowledge. Overall, these data challenge the integrity of semantic-conceptual development in DS and imply that contemporary theories of semantic cognition should also seek to incorporate evidence from atypical conceptual development. (181 words)

Keywords: semantic cognition, specific language impairment, Down syndrome, semantic, vocabulary
Receptive Vocabulary and Semantic Knowledge in Children with SLI and Children with Down Syndrome

Semantic memory, the long-term representation of meaning, underpins many language and cognitive functions. Contemporary research evidence about the location and organisation of semantic memory largely comes from studying adult neuropsychology patients, particularly those with semantic dementia (SD) who present with progressive loss of conceptual knowledge (e.g., Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000; Lambon Ralph, Cipolotti, Manes, & Patterson, 2010; Lambon Ralph & Howard, 2000). This research suggests that semantic memory is represented bilaterally in anterior temporal lobe regions (Lambon Ralph, Sage, Jones, & Mayberry, 2010; Pobric, Jefferies, & Lambon Ralph, 2007). The profile of patients with SD is characterized through deficits in word to picture matching, picture naming, category fluency, synonym generation, and tests of semantic association, relative to age- and education-matched controls (Bozeat et al., 2000). Bozeat et al. (2000) interpreted these results as an indication that the deterioration of semantic memory affected central conceptual knowledge that was independent of input modality. The sensitivity of their picture-based association test to degraded semantic representations suggests that a similar test might reveal weaknesses in children’s developing semantic systems. The present study adapted the procedure to investigate the vocabulary depths of children with specific language impairment (SLI), children with Down syndrome (DS) and a comparison group of typically developing children.

Semantic knowledge supports multiple aspects of typically developing children’s cognitive and language development, including short term memory (Bjorklund, 1987; Melby-Lervåg & Hulme 2010) and reading (Ouellette, 2006). Bjorklund (1987) suggests that as children acquire more detailed semantic knowledge, information becomes relatively more easily activated, providing an increase in processing efficiency that, in turn, facilitates
the use of memory strategies. Study of the semantic knowledge of very young typically developing children or children with speech or oral language impairments is limited by the availability of suitable assessments. Receptive vocabulary is generally assessed using standardised tests of word-picture matching such as the British Picture Vocabulary Scale II (BPVS II; Dunn, Dunn, Whetton, & Burley, 1997) or the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981). Such tests provide an indication of vocabulary breadth, the number of phonological entries within the lexicon that can be mapped to the correct semantic representations. However, vocabulary knowledge develops along a continuum from no knowledge, to partial or superficial understanding of words in relation to concepts, and eventually to the more nuanced understanding necessary for providing complete definitions or synonyms (Christ, 2011). Correct responses to a word-picture test of vocabulary demand a relatively superficial grasp of concept knowledge and give no indication of vocabulary depth, the extent to which word meanings have been refined and semantic knowledge has been elaborated (Christ, 2011; Funnell, Hughes, & Woodcock, 2006).

Contemporary models of concept learning draw on typicality and frequency of features with a parallel distributed processing simulation of activation across a learning network (McClelland & Rogers, 2003) arguably capturing the developmental shift from basic-level naming and feature awareness to more dispersed and fine-tuned knowledge of concepts at both superordinate and supraordinate levels. By simulating perturbations in the degree of activation of specific feature knowledge, this model has sought to explain progressive deterioration of semantic knowledge seen in SD (McClelland & Rogers, 2003). Developmental patterns of disrupted acquisition of concept knowledge are not well documented, perhaps due to a presumption of generalised, rather than specific, learning difficulties as a characterising feature of poor conceptual development.

Children with Specific Language Impairment
Children with SLI have language impairments that cannot be fully explained by intellectual disability, neurological or psychiatric diagnoses, or perceptual deficits (Bishop & Leonard, 2000; Leonard, 1998). Delayed lexical development is often an early sign of SLI (Bishop, 1997) but the diagnosis more often applies to children with a wider range of language and cognitive difficulties. Typically, phonology and grammar are more severely impaired than vocabulary although, even when the impairment predominantly affects expressive language, some children have delayed receptive vocabulary (Bishop, 1979, 1997).

One account of the SLI profile posits a deficit of the procedural learning system (Ullman & Piermont, 2005) that leaves the declarative memory system supporting the acquisition of semantic knowledge relatively intact. Early evidence suggested rapid incidental learning of new words after limited exposure can occur in children with SLI (Carey & Bartlett, 1978; Rice, 1990). These children readily learn the referents of new words despite being less able to retrieve the names of novel objects (Dollaghan, 1987). This implies that conceptual development and word-to-world mapping are positive aspects of the SLI profile, consistent with claims of the procedural deficit hypothesis (Ullman & Piermont, 2005). Interestingly, the availability of conceptual knowledge could provide compensatory support for weak phonological processing in children with SLI (following a neuroconstructivist perspective on cognitive development, Karmiloff-Smith, 1998) leading to a semantic strength hypothesis of SLI. Emerging evidence of overpriming in a pair-priming task (Pizziolo & Schelstraete, 2011) and semantic support for sentence processing tasks (Marinis & Van der Lely, 2007) provide some limited support for this hypothesis.

One challenge to a semantic strength hypothesis is to explain how word retrieval difficulties can occur in parallel (Pizzioli & Schelstraete, 2007). From the perspective of a parallel distributed processing theory of semantic cognition (Rogers & McClelland, 2004), strong multi-modal binding of features into concept knowledge should benefit all aspects of
word comprehension and production; although word finding difficulties could arise through atypical output in spoken word production and therefore be outside the scope of the model. A further challenge to the semantic strength hypothesis is evidence of weakly specified conceptual knowledge in children with SLI (McGregor, 1997; McGregor, Newman, Reilly, & Capone, 2002; Sheng & McGregor, 2010). Vocabulary depth was deficient when assessed through spoken semantic associates (Sheng & McGregor, 2010) and on speeded repetition tasks (Hennessey, Leitão, & Mucciarone, 2010). However, McGregor et al. (2002) found that five-to-eight-year-olds with SLI produced better drawings, gave more information about the attributes, and more accurate definitions, of objects named correctly than of misnamed objects or those not known. Interestingly, the evidence implies that conceptual knowledge accessed through word production tasks has less breadth for children with SLI, but that conceptual knowledge for ‘nameable’ (and retrievable) exemplars is enriched and readily activated within the language system.

**Individuals with Down syndrome**

Down syndrome (DS), or trisomy 21, is the most common biological cause of intellectual disability and affects around 1 in 1000 live births in the UK (Morris & Alberman, 2009). DS is associated with moderate to severe cognitive impairment with IQs in the range of 30 to 70 (Chapman & Hesketh, 2000; Glenn & Cunningham, 2005). Most individuals have language difficulties but there is wide variation in the severity of impairments and a characteristic profile of strengths and weaknesses with receptive vocabulary widely recognised as a developmental strength (e.g., Abbeduto, Warren, & Conners, 2007; Chapman, 1997; Chapman & Hesketh, 2000; Fowler, 1995; Gunn & Crombie, 1996; Laws & Bishop, 2004; Martin, Klusek, Estigarriba, & Roberts, 2009; Roberts, Price, & Malkin, 2007; Rondal, 1995; Rosin, Swift, Bless, & Vetter, 1988; Tager-Flusberg, 1999; Vicari, Caselli, & Tonucci, 2000; Ypsilanti & Grouios, 2008).
Evidence for this profile emerges in early childhood (Miller, 1999) with receptive vocabulary continuing to develop either in line with or in advance of nonverbal mental age (MA; Chapman, 1995; Fowler, 1995; Glenn & Cunningham, 2005; Laws & Bishop, 2003; Miller, 1999; Rosin, et al., 1988; Vicari, et al., 2000) and better than predicted either by mean length of utterance (Harris, 1983) or by syntactic comprehension (e.g., Chapman, Schwartz, & Kay-Raining Bird, 1991; Fowler, 1990, 1995; Gunn & Crombie, 1996; Laws & Bishop, 2003; Rosin, et al., 1988; Tager-Flusberg, 1999).

Fast mapping experiments suggest that individuals with DS learn new vocabulary as readily as MA-matched typically developing children (Chapman, Kay-Raining Bird, & Schwartz, 1990; Jarrold, Thorn, & Stephens, 2009; Kay-Raining Bird, Chapman, & Schwartz, 2004; Kay-Raining Bird, Gaskell, Babineau, & MacDonald, 2000). However, Jarrold et al. (2009) found that although individuals with DS did not differ significantly from younger typically developing children in learning the physical referents of novel words, they were impaired in learning the phonological forms, even when the task required no spoken responses. In this respect, fast mapping performance reflects the profile of children with SLI (Dollaghan, 1987) and the semantic strength hypothesis might equally apply to individuals with DS. A direct comparison of language abilities in DS and typical SLI, based on scores indexed by nonverbal MA and standardized relative to the scores of MA-matched typically developing children, showed similar profiles of strengths and weaknesses (Laws & Bishop, 2003). Despite the agreement among reviewers on the relative strength in semantic knowledge indexed by vocabulary breadth (e.g., Abbeduto, et al., 2007; Laws & Bishop, 2004), there has been little interest in understanding whether this is commensurate with vocabulary depth, possibly because of assessment difficulties.

Assessment of vocabulary depth has generally relied on children’s spoken definitions of words or attempts to explain how two things are alike (e.g., Ouellette, 2006; Ouellette &
Beers, 2010; Wechsler, 1999), or on interviews to probe children’s deeper understanding of word meanings (e.g., Funnell et al., 2006). Such assessments may provide unreliable access to the semantic knowledge of children with language impairments due to the confounding effects of word finding or other expressive language difficulties. Speech production difficulties, experienced by many children with DS (Cleland, Wood, Hardcastle, Wishart, & Timmins, 2010) and some children with SLI (Shriberg, Tomblin, & McSweeney, 1999) may also limit responses. Semantic priming experiments (e.g., Nation & Snowling, 1999) offer alternative, implicit assessments of semantic association, although assessments based on printed words are inappropriate for some children with SLI or with DS with more limited reading skills (e.g., Catts, Bridges, Little, & Tomblin, 2008; Laws, 2010). Auditory lexical decision tasks can reveal the strength of semantic associations (e.g., Pizzioli & Schelstrate, 2011) but performance could be affected by hearing loss, which is common in DS (Davies, 1996), or by severely limited memory spans that, in some individuals with DS, do not extend to two items (e.g., Laws & Gunn, 2004).

Despite many descriptions of receptive vocabulary strength, little is known about the depth of vocabulary knowledge of individuals with DS, or in developmental disorders such as SLI, or about how vocabulary breadth and depth are related in these conditions.

Outline of study

Vocabulary breadth and depth were assessed in children with SLI or with DS and results were compared to those of typically developing children matched on raw scores from the BPVS II (Dunn, et al., 1997) as a word-picture matching task. To investigate vocabulary depth, a task was devised based on tests used to assess semantic association in adults with SD (Bozeat et al., 2000; Howard & Patterson, 1992). In the Camel and Cactus Test (Bozeat, et al., 2000), participants see a picture (for example, a camel) and select its associate from an array of four other pictures (in this example: cactus [correct response], tree, sunflower or
A version with items suitable for children was developed, which we refer to as the Baby and Pram Test. Items for the test were drawn from the receptive vocabulary test, allowing direct comparison of vocabulary breadth with vocabulary depth. Vocabulary breadth and depth were compared within each group and across the three study groups.

The main research question was whether the language-impaired groups would have concept knowledge comparable to that of typically developing children with similar breadths of vocabulary. A semantic strength hypothesis would predict that children with SLI should have greater conceptual knowledge indexed through vocabulary depth than the typical children. However, other research, notably that of McGregor and colleagues, would predict weaker vocabulary breadths than typical children (that is, children with SLI would be older than the vocabulary-matched children, despite the similarity in their vocabulary scores) and show commensurate vocabulary depth. Despite well-established research support for receptive vocabulary strength in DS, there is no basis for predicting how vocabulary depth will be related to vocabulary breadth. The similarity with the language profile associated with SLI (e.g., Laws & Bishop, 2003) could suggest similar profiles across word-picture matching and picture association tasks, especially when identical lexical items are employed, based on the assumption that concept formation is relatively intact in both groups, even if the rate of acquisition of concepts is slower than for typically developing children.

Method

Participants

All the children attended mainstream schools or nurseries. Parents confirmed that children spoke English as their first language, had no history of hearing difficulties and had received no diagnoses of autism or other developmental disorders (other than SLI or DS).
Sixteen children (9 boys, 7 girls, $M_{\text{age}} = 7$ years, 7 months, range = 6 years, 5 months to 9 years) were selected from a group taking part in a longitudinal project to investigate oral language and reading development. Eleven children were recruited from two language units attached to mainstream schools, four were referred by speech and language therapists, and one child was recruited through another school. Children had received diagnoses of SLI from speech and language therapists, had language difficulties consistent with this diagnosis according to the Children’s Communication Checklist, version 2 (Bishop, 2003), scored 1 standard deviation or more below the mean on 2 to 5 of 5 oral language tests, and had nonverbal IQs > 80 on recruitment to the main project. To ensure meaningful comparison, children were selected for the current study on the basis that receptive vocabulary scores were within the range obtained by the children with DS (described below). This resulted in the exclusion of three children with scores that were higher than those recorded for any child with DS. All the children passed a hearing screen with maximum thresholds of 25 dB measured at frequencies of 500, 1000, 2000 and 4000 Hz.

**Children with Down Syndrome.**

Sixteen children (6 boys, 10 girls, $M_{\text{age}} = 10$ years, 2 months, range = 6 years 10 months to 13 years) were selected from the same longitudinal study (Laws, 2010). Most children were recruited through parent support groups, four through education services, and one through speech and language therapy services. Only children with receptive vocabulary scores within the range achieved by children with SLI (see above) were selected. This resulted in the exclusion of twelve children with scores that were lower than those recorded for any child with SLI. Most included children were reported to have trisomy 21; one girl had mosaicism but, since her vocabulary and nonverbal cognitive abilities were within the ranges obtained for the sample, she was retained in the study. Good hearing was not an inclusion criterion but, although hearing loss is common in children with DS, none of the
participants had hearing difficulties reported either by parents or by the children’s audiologists.

**Typically developing children.**

Nineteen typically developing children (12 boys, 7 girls, $M_{\text{age}} = 4$ years, 11 months, range = 3 years, 5 months to 7 years, 6 months) provided a comparison group. Participants were selected from a larger group recruited from mainstream primary schools on the basis that they had no known educational or learning difficulties and had receptive vocabulary scores within the normal range for CA: BPVS II standard scores ranged from 92 to 123 ($M = 103.4$, $SD = 7.85$).

**Group matching procedures**

We aimed to match each child from the clinical groups to a typical child on receptive vocabulary scores. In the group with DS, 13/16 children were so matched to within -1 and +1 points; 3 children were less closely matched with scores between 3 and 5 points higher than the child with whom they were paired. In the group with SLI, 11 children were matched to the same typically developing children to within -1 and +1 points with 5 children matched to within -2 and +3 points. Where two potential typical matches were available, both children were included in the comparison group.

One-way ANOVA showed no significant effect of group on BPVS II raw scores, $F(2, 48) = .349$, $p = .71$. Following Kover and Atwood (2013), adequacies of matches between groups were assessed by inspecting $p$-values for differences between means, effect sizes (Cohen’s $d$), and variance ratios, aiming for $p$-values $\geq .5$, Cohen’s $d$ close to 0, and variance ratios close to 1 (see Table 1).

***Table 1 about here***

The SLI/TD group comparison closely approximated these values. The values for comparisons involving the group with DS were not so close. However, since thresholds for
determining adequacy should be considered in light of the clinical implications of group
differences and the balance of the effect size of the matching variable relative to the effect
size for the dependent variable (Kover & Atwood, 2013), these matches were considered
satisfactory.

**Preliminary description of groups by chronological age**

One-way analysis of variance (ANOVA) confirmed a significant between-groups effect
of CA, $F(2, 48) = 59.71, p < .0001$. Post-hoc tests using the method of Least Significant
Difference showed that the group with DS was significantly older than the group with SLI
and that both groups with language impairments were significantly older than the typically
developing comparison group ($p$s < .0001).

**Assessment and Measures**

**Nonverbal cognitive ability.**

Children with DS or SLI were assessed using the Leiter International Performance
Scale – Revised (Leiter-R; Roid & Miller, 1997). Following directions in the manual, scores
on four subtests (figure ground; form completion; sequential order; and repeated patterns)
were used to calculate nonverbal mental ages (Leiter MAs).

**Receptive vocabulary.**

The BPVS II (Dunn et al., 1997) measured receptive vocabulary. For each item, the
examiner speaks a word and asks the child to choose the picture corresponding to that word
from a display of four pictures. Children were assessed from the beginning of the test rather
than from the starting points for CAs indicated in the test instructions.

**Vocabulary breadth**

The BPVS II test presents items in sets of 12 with successive sets representing
increasing difficulty. 35 items from the test were selected to provide a measure of
vocabulary breadth. Only concrete nouns were selected, avoiding verbs and adjectives. Most
items came from Sets 1 (e.g., baby) to 5 (e.g., hive) but one item was selected from each of Sets 6 and 8, extending task difficulty.

Vocabulary depth.

The same 35 nouns were used as items for the Baby and Pram Test as a measure of vocabulary depth. For each item (e.g., baby), children were offered four response picture choices (in this example: pram\(^1\) [correct response], wheelbarrow, go-kart, bicycle). Correct responses were semantically related to items and to foils but the foils were not related to the items. The majority of associations depended on function (e.g., baby and pram). Nine items depended on analogy (e.g., tortoise and snail); one item depended on shared category membership (eagle and owl); and three items depended on general knowledge (e.g., arrow and Robin Hood). Coloured photographs of the items and response pictures were sourced from on-line and CD photograph libraries. Pictures were independently validated by three researchers. Two items were replaced for the final version after conducting a pilot study with six typically developing children aged 7 to 8 years.

The test was programmed using E-Prime, version 1.0 (Psychology Software Tools, Inc.) and presented on a 15-inch Elo USB Touchscreen controlled by a laptop computer. Items were arranged randomly into four fixed lists and participants received the lists in one of four orders. Participants sat facing the screen; the target picture appeared at the top and smaller pictures of the four response choices appeared simultaneously in a row beneath. The position of the correct response picture was varied across items. Children were asked to review the response choices and to touch the one that went with the bigger picture at the top. The program recorded response accuracy and reaction times but no emphasis was placed on quick responses and children were encouraged to think carefully before making their selection. To ensure that task demands were understood, two easy practice trials were offered (car and

\(^1\) Stroller in the United States
wheel; table and chair). All the participants demonstrated that they understood the task before proceeding.

**Procedures.**

The Faculty of Science Research Ethics Committee at the University of Bristol and Gloucestershire NHS Research Ethics Committee approved the research. All parents were sent information and returned a signed consent form before children were assessed in school or nursery. For the clinical groups and the older typically developing children, some measures formed part of a battery of assessments administered for the main project. The youngest typically developing children (under five years) were assessed at a later date and completed only the vocabulary measures.

**Results**

Table 2 describes chronological age and test scores for the three groups.

***Table 2 about here***

BPVS II raw scores were converted to age-equivalent scores that indicate the age at which a typically developing child would be expected to achieve a given score, which we refer to as vocabulary age (VA). The group with DS had mean VA significantly below the level expected for CA, mean difference = 55 months, \( t(15) = 10.563, p < .0001, 95\% \ CI [43.55, 65.57]\). The group with SLI also had mean VA significantly lower than expected for CA, mean difference = 20 months, \( t(15) = 8.721, p < .0001, 95\% \ CI [15.44, 25.43]\). Thus, vocabulary was delayed relative to CA in both clinical groups.

For the group with DS, mean VA was significantly in advance of Leiter MA, \( t(15) = 4.10, p < .0001, 95\% \ CI [5.64, 17.86]\), with a mean difference of 11.75 months. The same comparison for the group with SLI showed that VA was significantly below Leiter MA, \( t(15) = -4.01, p < .001, 95\% \ CI [-18.00, -5.50]\), with a mean difference of 11.75 months in the
opposite direction. In summary, the receptive vocabulary of children with SLI was lower than would be predicted either by CA or nonverbal MA, while the group with DS demonstrated strength in receptive vocabulary relative to nonverbal MA.

**Vocabulary Breadth and Depth**

For each child, vocabulary breadth was calculated as the number of items correctly identified from the subset of BPVS II items on which the Baby and Pram Test was based. Vocabulary depth was calculated as the number of items correctly identified on the Baby and Pram Test (see Table 3). Cronbach’s alphas for vocabulary breadth and depth tests were .7 and .9 respectively.

****Table 3 about here****

For each participant, responses to each item were classified as: (i) incorrect responses on both tests; (ii) item correctly recognised on the vocabulary breadth test but its semantic associate was not known; (iii) the correct semantic associate was chosen but the item itself was not identified; and (iv), correct responses were provided on both tests. For each group, the mean number of items with each pattern of response was calculated (see Table 4).

***Table 4 about here***

Nearly all children scored more on the test of vocabulary breadth than on the test of vocabulary depth; scores were equal for 1 child with DS and 2 typical children, and 1 child with SLI scored one point more on the breadth test. In all three groups, it was rare for a child to correctly identify the semantic associate of an item on the Baby and Pram Test that they could not also point to on the receptive vocabulary test.

**Pattern of vocabulary breadth and depth scores for typically developing children.**

For the typically developing group, mean vocabulary breadth was significantly higher than mean depth, mean difference = -5.37, t (19) = -5.83, p < .0001, 95% CI [-7.30, -3.43].
To ascertain whether there were differences between older and younger children, as in Funnell et al.’s (2006) study, results were investigated separately for younger children, aged up to six years, five months (n = 12), and older children, aged six years, six months and over (n = 7). A mixed ANOVA confirmed the significant effect of task, $F(1, 17) = 28.66, p < .0001, \eta^2_p = .63$. A significant between-groups effect, $F(1, 17) = 8.94, p = .008, \eta^2_p = .35$ was explained by overall higher scores for the older group. However, critically, there was no significant task by group interaction, $F(1, 17) < 1$, and so no evidence that older children responded to the tasks differently from younger children. Typically developing children of all ages tended to do better on the test of vocabulary breadth than that of vocabulary depth.

**Vocabulary breadth and depth of the groups with DS and SLI compared to typically developing children.**

To investigate the patterns of results for the groups with language impairments, taking into account the differences in performance across the tasks evident for typically developing children, vocabulary breadth and depth scores for children with DS and with SLI were standardized relative to the performance of the typically developing group (see Figure 1).

***Figure 1 about here***

The resulting z-scores were used in a mixed ANOVA treating standardized vocabulary breadth and depth scores as repeated measures and study group (TD v SLI v DS) as the between-groups measure. The between-groups effect was significant, $F(2, 48) = 12.45, p < .0001, \eta^2_p = .34$. Pairwise comparisons indicated that DS group z-scores across tasks differed significantly from TD means of 0, mean difference = -1.46, $p < .0001$, 95% CI [-2.13, -.78], and also differed significantly from SLI group z-scores, mean difference = -1.48, $p < .0001$, 95% CI [-2.26, -.70]. However, SLI group z-scores across tasks did not differ significantly from TD means of 0, mean difference = .03, $p = .92$, 95% CI [-.56, .62].
A trend towards a significant effect of task, $F (2, 48) = 3.99$, $p = .051$, $\eta^2_p = .07$, reflected the fact that, across groups, mean breadth $z$-scores were higher than mean depth $z$-scores (breadth: $M (SD) = -.26 (1.11)$; depth: $M (SD) = -.64 (1.70)$). However, this was qualified by a significant interaction between task and group, $F (2, 49) = 9.04$, $p < .0001$, $\eta^2_p = .27$. Separate one-way ANOVAs to investigate differences in group means for each of the vocabulary measures indicated that there was no significant group effect on breadth $z$-scores, $F (2, 49) = 1.58$, $p = .22$, but group membership did significantly influence depth $z$-scores, $F (2, 49) = 18.09$, $p < .0001$. Post-hoc tests confirmed that the group with SLI did not differ significantly from the TD group. The group with DS scored significantly lower on the vocabulary depth test than the TD group, mean difference = -2.26, $p < .0001$, 95% CI [-3.22, -1.31], and significantly lower than the group with SLI, mean difference = -2.50, $p < .0001$, 95% CI [-3.55, -1.44].

Within the SLI group, there was no significant difference between the standardised breadth and depth scores, breadth: $M = -.18$, $SD = 1.06$; depth: $M = .23$, $SD = 1.013$, $t (15) = 1.18$, $p > .05$, 95% CI [-.97, -.12]. However, within the group with DS, standardised breadth scores were significantly greater than standardised depth scores, breadth: $M = -.65$, $SD = 1.24$; depth: $M = -2.26$, $SD = 1.74$, $t (15) = -3.50$, $p = .003$, 95% CI [.63, 2.6]. In summary, children with DS with receptive vocabulary in line with that of children with SLI and of typically developing children did not have commensurate depths of semantic knowledge.

**Correlations among study measures.**

Vocabulary breadth and depth were significantly correlated within the group with SLI, $r (16) = .54$, $p = .01$. However, there was no significant correlation between the measures in the group with DS, $r (16) = .27$, $p = .158$ or with TD, $r (19) = .15$, $p = .27$. Table 4 shows concurrent correlations between vocabulary breadth and depth with CA and Leiter MA within each group.
After adjusting accepted probability to $p < .005$ to take account of multiple correlations, depth but not breadth of vocabulary was significantly correlated with CA within all three groups. Other moderate correlations amongst the measures did not reach statistical significance ($ps = .01 - .04$).

**Discussion**

This study investigated vocabulary breadth and depth in children with SLI, children with DS and typically developing children. The Baby and Pram Test was devised to be less challenging for children with language impairments as a measure of vocabulary depth than verbal assessments such as the elicitation of definitions or synonyms (e.g., Ouellette, 2006). The test provided a reliable measure of vocabulary depth and the pattern of results across the two measures supported previous findings that children can identify the names of objects before elaborating their semantic knowledge of them.

**Receptive Vocabulary in DS and in SLI**

The receptive vocabulary of children with SLI was weak relative to CA and to Leiter MA. Although only 5/16 children with SLI scored below a level that would generally be considered to represent impairment (BPVS II standard scores below 85), all had VAs below CAs and 13/16 children also had VAs below Leiter MAs, confirming a picture of slow acquisition of word knowledge consistent with that reported in previous research (e.g., McGregor, et al., 2002).

The results for the children with DS were also consistent with previous research, confirming a developmental strength in receptive vocabulary relative to nonverbal MA (e.g., Chapman, 1995; Glenn & Cunningham, 2005; Laws & Bishop, 2003). Mean VA was significantly in advance of mean Leiter MA, a pattern of results that was observed for 14/16 children with DS. Glenn and Cunningham (2005) suggested that BPVS II and Leiter-R tests...
are interchangeable for the purpose of establishing the MAs of individuals with DS. However, the lack of correlation between the measures in our experiment cautions against this advice. VA was in advance of Leiter MA for most children but there was wide variation in the differences between the measures, ranging from 1 to 39 months.

**Vocabulary Breadth and Depth in children with SLI**

The main research question was whether children with language impairments would resemble typically developing children in terms of associated semantic knowledge. The standardized mean breadth and depth vocabulary scores of the group with SLI did not differ significantly from the typically developing comparison group means. In both groups, children achieved higher scores for the measure of vocabulary breadth than for vocabulary depth. This pattern of results is consistent with the idea that children with SLI also acquire word forms before elaborating word meanings, and support Bishop’s (1997) conclusion that the causes of slow lexical development in SLI do not encompass difficulties with the acquisition and retrieval of semantic knowledge. Although mean vocabulary depth scores were higher than the mean for the (younger) typically developing children, the difference was not statistically significant and did not support the semantic strength hypothesis.

**Deficit and Dysregulation in Semantic Knowledge of Children with DS**

The results provide evidence for a significant impairment in acquiring concept knowledge for the group with DS. The dissociation between vocabulary breadth and depth was not due to difficulties in understanding the task demands of the Baby and Pram Test since children succeeded on the easiest items. To some extent this pattern of results was surprising since children with DS could have been predicted to find the Baby and Pram Test easier than the word-picture matching test because, first, the Baby and Pram is entirely picture-based and second, children do not actually need to reproduce the names of the items to succeed. For example, a picture of a penguin could be correctly associated with a picture
of a fish without it being necessary to generate the names of either creature. The word-picture matching test should be more difficult because the child must attend to the word *penguin* and then integrate his or her phonological representation of the word with its meaning.

In order to provide comparable samples, only children with DS with higher vocabulary abilities had been selected for study. This raises a question about how far the findings can be generalised to the wider population with DS and suggests the need for further research to learn more about the semantic knowledge of individuals with DS with lower verbal abilities. As in the other two groups, Baby and Pram Test scores increased with CA in the group with DS. Two older children (aged 13 years and 11 years 7 months) achieved test scores that were comparable to their vocabulary breadth scores (28 and 29/35 and 31 and 31/35, respectively). Since the acquisition of semantic knowledge depends on life experience, it is not surprising to find that older individuals achieved higher scores overall. However, further research should establish whether semantic deficits generally diminish with chronological age in individuals with DS, or whether it is more usual for deficits in semantic knowledge to persist.

One explanation of the deficit draws on a parallel distributed processing model of semantic cognition (Rogers & McClelland, 2004). Under this account, conceptual development in DS is impaired, not simply through weaker patterns of distributed activation of modality-specific input that contribute to delayed pattern of acquisition, but additionally through weak convergent activation in the anterior temporal lobes that generates multi-modal concept knowledge; a *conceptual deficit account*. However, the neuroanatomical evidence for this type of characterisation of DS is scarce; individuals with DS typically demonstrate weaker hippocampal volumes as adults and children (Pinter, Eliez, Schmitt, Capone & Reiss, 2001), rather than abnormalities of the anterior portion of the temporal lobe. Behavioural
An alternative possibility is that active consideration of multiple visual distractors in the Baby and Pram Test places demands on strategic and controlled processes that influenced children’s decision-making. For any given trial, multiple semantic features are available that potentially increase the inhibitory demands of suppressing semantic / featural information and selecting more proximal associations for a positive decision towards the correct item. Jefferies and Lambon-Ralph (2006) highlighted a role for executive processes in “directing and shaping linguistic and semantic activation” that has been further identified through a profile of dysregulation of semantic knowledge in patients with transcortical sensory (stroke) aphasia (Hoffman, Jefferies, & Lambon-Ralph, 2011). In this way, a semantic ‘dysregulation’ account could contribute to the marked impairment of the individuals with DS and could align with their broader profile of difficulties in nonverbal cognition.

Implications of Dissociation of Receptive Vocabulary from Semantic Knowledge in DS

Glen and Cunningham (2005) suggested that cross-syndrome comparisons of functions that depend on verbal abilities should employ BPVS II as a matching measure. However, the dissociation between vocabulary breadth and depth implies that individuals with DS matched on receptive vocabulary to other clinical or to typically developing groups will differ from them in ways that could potentially influence investigations of other functions. Although receptive vocabulary scores were equivalent to those of typical five-to-seven-year-olds, the level of semantic knowledge evident for children with DS was at or below that of typical four-year-olds. These results support the idea that lexical-semantic development in DS is not simply a slower version of typical development but follows a fundamentally different course that could be reflected in a variety of different paradigms including short-term memory and syntactic measures; for example, semantic errors in the generation of relative clauses are
more likely for individuals with DS than those with SLI (Stathopoulou & Clahsen, 2007). Although the neuroanatomical basis for impaired semantic cognition in DS is poorly specified, these data contribute to emerging evidence of specific impairments to semantic cognition to be identified in developmental populations (Briscoe, Chilvers, Baldeweg, and Skuse, 2012; Nation and Snowling, 1999).

The dissociation between vocabulary breadth and depth has implications for teaching and for speech and language therapy. Evidence for impoverished semantic knowledge suggests that interventions to deepen vocabulary knowledge could be worthwhile for individuals with DS. A recent intervention involving children with DS included an activity to build ‘word-webs’ to reinforce understanding by linking words to associated concepts (Burgoyne, Duff, Clarke, Buckley, Snowling, & Hulme, 2012). A poor response to direct teaching of receptive vocabulary, even after an intensive, 20 week intervention, provides further evidence of an atypical developmental pathway for the acquisition of conceptual knowledge in individuals with DS. Further research should establish the nature of these differences with a view to developing new approaches to teaching and therapy.

**Summary and Conclusions**

In conclusion, the Baby and Pram Test provided a reliable measure of children’s semantic knowledge. The results indicated that typically developing children and children with SLI acquire word-picture vocabulary in advance of broader conceptual knowledge. Although the rate of vocabulary acquisition is slowed, the depth of vocabulary knowledge of children with SLI was commensurate with vocabulary breadth. In contrast, despite strength in receptive vocabulary relative to nonverbal MA, children with DS demonstrated significant deficits in semantic knowledge. More research is necessary to understand the theoretical origin of semantic deficits and to establish whether improving the extent and structure of children’s knowledge bases would generalise to improvements in other functions.
Acknowledgements

The authors thank the children who participated, and the teachers and schools that cooperated with the research.

References


Table 1. Descriptive statistics for BPVS II raw scores used to match groups with TD, SLI and DS

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD (n = 19)</td>
<td>59.84</td>
<td>8.81</td>
<td>77.59</td>
</tr>
<tr>
<td>SLI (n = 16)</td>
<td>60.00</td>
<td>8.52</td>
<td>72.67</td>
</tr>
<tr>
<td>DS (n = 16)</td>
<td>57.69</td>
<td>9.19</td>
<td>84.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Difference</th>
<th>p</th>
<th>Cohen’s d</th>
<th>Variance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLI/TD</td>
<td>0.16</td>
<td>.85</td>
<td>.02</td>
<td>.94</td>
</tr>
<tr>
<td>DS/TD</td>
<td>2.15</td>
<td>.49</td>
<td>.24</td>
<td>1.09</td>
</tr>
<tr>
<td>DS/SLI</td>
<td>2.31</td>
<td>.47</td>
<td>.26</td>
<td>1.16</td>
</tr>
</tbody>
</table>

*Note. TD = typical development; SLI = specific language impairment; DS = Down syndrome*
Table 2. Description of receptive vocabulary and nonverbal cognitive abilities of groups with TD, SLI and DS.

<table>
<thead>
<tr>
<th>Variable</th>
<th>TD (n = 19)</th>
<th>SLI (n = 16)</th>
<th>DS (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA $^a$</td>
<td>5;7 (1;2)</td>
<td>7;7 (0;8)</td>
<td>10;2 (1;7)</td>
</tr>
<tr>
<td>Leiter</td>
<td>6;10 (1;2)</td>
<td>4;8 (0;7)</td>
<td></td>
</tr>
<tr>
<td>MA$^{a,b}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BPVS II rs</td>
<td>59.84 (8.81)</td>
<td>60.00 (8.52)</td>
<td>57.69 (9.19)</td>
</tr>
<tr>
<td>BPVS II ss</td>
<td>103.63 (7.99)</td>
<td>86.19 (6.00)</td>
<td>66.00 (11.68)</td>
</tr>
<tr>
<td>BPVS II</td>
<td>5;10 (0;11)</td>
<td>5;10 (0;10)</td>
<td>5;8 (0;11)</td>
</tr>
</tbody>
</table>

$^a$ years; months. $^b$ Not completed by children with TD.

Note. TD = typical development; SLI = specific language impairment; DS = Down syndrome; CA = chronological age; MA = mental age; BPVS II = British Picture Vocabulary Scales II; rs = raw score; ss = standard score; VA = vocabulary age.
Table 3. Mean (SD) vocabulary breadth and vocabulary depth scores for groups with TD, SLI and DS.

<table>
<thead>
<tr>
<th>Variable</th>
<th>TD (n = 19)</th>
<th>SLI (n = 16)</th>
<th>DS (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocab breadth (max = 35)</td>
<td>29.68 (2.11)</td>
<td>29.31 (2.24)</td>
<td>28.31 (2.63)</td>
</tr>
<tr>
<td>Vocab depth (max = 35)</td>
<td>24.32 (3.76)</td>
<td>25.19 (4.25)</td>
<td>15.81 (6.52)</td>
</tr>
</tbody>
</table>

Note. TD = typical development; SLI = specific language impairment; DS = Down syndrome; max = maximum; vocab = vocabulary.
Table 4. Scoring patterns across breadth and depth items: mean (SD) number of items with each response pattern for groups with TD, SLI and DS

<table>
<thead>
<tr>
<th>Scoring pattern</th>
<th>TD (n = 19)</th>
<th>SLI (n = 16)</th>
<th>DS (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect on both tests</td>
<td>1.4 (1.1)</td>
<td>1.8 (1.8)</td>
<td>3.7 (2.1)</td>
</tr>
<tr>
<td>Breadth only correct</td>
<td>9.3 (3.3)</td>
<td>6.2 (4.1)</td>
<td>15.1 (5.3)</td>
</tr>
<tr>
<td>Depth only correct</td>
<td>2.8 (1.2)</td>
<td>2.9 (1.9)</td>
<td>2.4 (1.5)</td>
</tr>
<tr>
<td>Both correct</td>
<td>20.5 (3.8)</td>
<td>23.1 (5.3)</td>
<td>13.1 (6.0)</td>
</tr>
</tbody>
</table>

Note: TD = typical development; SLI = specific language impairment; DS = Down syndrome.
Table 5. Concurrent correlations of breadth and depth of vocabulary scores with CA and Leiter MA.

<table>
<thead>
<tr>
<th></th>
<th>TD (N = 19)</th>
<th>SLI (N = 16)</th>
<th>DS (N = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vocabulary breadth</td>
<td>Vocabulary depth</td>
<td>Vocabulary breadth</td>
</tr>
<tr>
<td>CA</td>
<td>.46</td>
<td>.65*</td>
<td>.42</td>
</tr>
<tr>
<td>Leiter MA</td>
<td>.44</td>
<td>.57</td>
<td>.54</td>
</tr>
</tbody>
</table>

*Note. TD = typical development; SLI = specific language impairment; DS = Down syndrome; CA = chronological age; MA = mental age

*p ≤ .003, one-tailed.
Figure labels

*Figure 1.* Mean z-scores for DS and SLI groups standardized in relation to vocabulary breadth and depth scores of typically developing children.