

# A LONGITUDINAL STUDY OF CHILDREN WITH DEVELOPMENTAL LANGUAGE DELAY AT AGE THREE: LATER INTELLIGENCE, READING AND BEHAVIOUR PROBLEMS

Phil A. Silva  
Sheila Williams  
Rob McGee

Stevenson (1984), in a discussion of research on the predictive value of speech and language screening, reviewed three studies which followed up children who had had language delays in the preschool period. The first was carried out in Newcastle (Fundudis *et al.* 1979), the second in Waltham Forest, London (Richman *et al.* 1983) and the third in Dunedin, New Zealand (Silva *et al.* 1983). All three studies identified their language-delayed samples and comparison groups when they were aged about three years and followed them up at the ages of seven (Newcastle and Dunedin) or eight (Waltham Forest). All three studies reported later IQ and reading-test scores. The Newcastle and Waltham Forest studies, but not the Dunedin study, also reported results from the Rutter Teacher Scale (Rutter *et al.* 1970).

Results from the three studies were strikingly similar. All found that children with early language delay had significantly more low IQs (particularly verbal IQs) and either a significantly lower mean reading score (Newcastle) or a significantly higher prevalence of poor readers in comparison to controls. The prevalence of behavioural problems among those children from Newcastle and Waltham Forest who had been language-delayed as preschoolers was very similar (22 and 24 per cent, respectively). Because of the

different methods of drawing the samples, comparison of the behavioural results was not possible; however, both studies showed a discernible tendency for the language-delayed children to have a higher degree of later behavioural deviance than children without such delay. Stevenson speculated that 'if the behavioural status of the Dunedin sample had been measured, it is likely that a similarly high percentage of children would have been identified' (p. 534).

This report extends those results by describing follow-up studies at ages seven, nine and 11 years of children who were language-delayed at age three. Intelligence and reading-test results are described, together with behaviour problems reported by teachers and parents.

## Method

### Sample

The sample was followed longitudinally by the Dunedin Multidisciplinary Health and Development Research Unit and has been described previously (Silva 1980, Silva *et al.* 1983). The study and sample are described more fully by McGee and Silva (1982). The sample was advantaged in terms of socio-economic status in comparison with the rest of the country, and under-representative of the Maori and other Polynesian races.

The original sample comprised all

children born at Dunedin's one obstetric hospital during the 12-month period between 1st April 1972 and 31st March 1973 who were known to be resident in Otago at age three years ( $N=1139$ ). Of these, 1037 (91 per cent) were followed up as three-year-olds. Those not followed up were either refusals ( $N=68$ ) or were located too late for inclusion ( $N=34$ ). The children were assessed as five-year-olds ( $N=991$  or 96 per cent of the original 1037), seven-year-olds ( $N=954$  or 92 per cent), nine-year-olds ( $N=955$  or 92 per cent) and 11-year-olds ( $N=925$  or 89 per cent). Of the 112 children who were not followed up at age 11, 70 were not assessed because of parental refusal to co-operate, 38 because they lived in areas where it was not possible to arrange assessment and four children had died. McGee (1984) reported a study to determine whether there were any significant differences between children who were or were not assessed at age 11 on measures taken in earlier phases of the research. The measures included tests of language development, intelligence and reading, and the Rutter Behaviour Scales. There were no significant differences between those assessed and those not assessed as 11-year-olds on any measure of child development. This suggests that those who dropped out did not bias the remaining sample.

#### *Data collection*

Most of the children were assessed within approximately one month of their third, fifth and seventh birthdays and within approximately two months of their ninth and 11th birthdays. Those assessed outside these age-ranges had their scores adjusted for age where appropriate. All assessments were carried out by trained psychometrists, without reference to any previous results.

#### *Measures*

The Reynell (1969) Developmental Language Scales were used to define language delay at age three. At ages seven, nine and 11, intelligence, reading and behaviour were assessed with the same measures. Intelligence was assessed using the Wechsler Intelligence Scale for Children (WISC-R; Wechsler 1974). To

save time, one subtest from each scale was omitted (Comprehension and Picture Arrangement). Results were prorated according to the test manual recommendations. Reading was assessed by the Burt Word Reading Test (Scottish Council for Research in Education 1976). Previous research had shown that results from this test correlate highly with those from a prose reading test ( $r=0.86$ ) (Silva 1981) and moderately highly with tests of reading comprehension (mean  $r=0.69$ ) and reading vocabulary (mean  $r=0.75$ ) (Gilmore *et al.* 1981). Behaviour was assessed with the Rutter Parent and Teacher Scales (Rutter *et al.* 1970). Some results from the use of these scales in the Dunedin study are described by McGee *et al.* (1983, 1984, 1985).

The children's backgrounds were assessed by use of a family disadvantage index, described by McGee *et al.* (1985), together with some correlates of this type of disadvantage. The index was derived from information obtained from the assessment at age five. One index point was given for each of the following: low socio-economic status (Elley and Irving 1972) at time of birth of the child; mother's age less than 21 at birth of her first child; separations from the child's father in the first five years, including being solo at birth; large family size (four or more children); whether the parents had consulted a marriage guidance counsellor; and a high score by the mother on a malaise inventory (Rutter *et al.* 1970).

#### *Definitions of language delay, low intelligence, low reading score and behaviour problems*

As with our previous study (Silva *et al.* 1983), a three-year-old was considered to have language delay if he or she scored at or below the nearest score on the Reynell scales that corresponded to the fifth percentile on either the verbal comprehension or verbal expression scales. The actual scores were 18 for verbal comprehension and 21 for verbal expression. These scores corresponded to the mean score of children 10 to 12 months younger than the three-year-olds (Reynell 1969). By this definition, children could be delayed in verbal

comprehension only, verbal expression only, or in both aspects ('general language delay').

A low IQ was defined as one of 89 or less, the point defined by Wechsler (1974) as the lower end of the range of normal intelligence.

Different low reading-test scores were defined for each age. At age seven the sample mean for six-year-olds (score 12) was taken as the cut-off point; at age nine that for seven-year-olds (30); and at age 11 that for eight-year-olds (42).

Behaviour problems were defined according to the recommendations of Rutter *et al.* (1970), *i.e.* children with a score  $\geq 13$  on the Rutter Parent (Form A) Scale and those with a score of  $\geq 9$  on the Teacher (Form B) Scale were defined as having a significant degree of problem behaviour.

#### Statistical Analysis

The SPSS-X (1983) procedure MANOVA was used to analyse the results. This procedure tests for a group-by-time interaction effect; that is, it tests whether the profiles of the means are parallel. If the interaction effect is not statistically significant, the procedure tests the null hypothesis that there are no differences among the groups. In this study, if the differences among the groups were statistically significant, pair-wise Scheffe *post hoc* tests ( $p < 0.05$ ) were carried out. The time effects were not of interest.

Chi-squared tests were used to analyse the categorical data. If these were significant they were followed by *post hoc* tests which compared each of the language-delay groups with the remainder of the sample (Everitt 1977).

#### Results

##### Identification of children with language delay at age three years

All but 10 of the 1037 three-year-olds were successfully assessed with the Reynell Scales. Of these, 78 (7.6 per cent) scored at or below the cut-off point on one or both scales: 27 (2.6 per cent) were delayed in verbal comprehension only, 24 (2.3 per cent) in verbal expression only, and 27 (2.6 per cent) in both aspects (general language delay). Each language-delay group had significantly more boys

than girls ( $p < 0.05$ ). There were 19 boys and eight girls in the comprehension-delayed group, 18 boys and six girls in the expressive-delayed group and 19 boys and eight girls in the group with general language delay. There were 474 boys and 475 girls in the remainder of the sample.

##### Descriptive results at ages seven, nine and 11

Table I shows the numbers of children, means and standard deviations for all measures used at ages seven, nine and 11 for the total sample, the cut-off points at or below which children were taken to have low intelligence or low reading scores and at or above which children were taken to have behaviour problems, and the numbers of children falling below, at or above the cut-off points.

The mean IQs for the sample ranged from 103.1 to 110.9. The mean reading scores increased with age. The mean Rutter Parent Behaviour Scale scores ranged from 7.2 to 8.8 and the Teacher Scale scores from 3.8 to 4.3. For both behaviour scales the 11-year means were slightly lower than the seven-year means. Between 8 and 23 per cent of the children fell at or below the cut-off points for low intelligence or low reading scores or at or above the cut-off points for having behaviour problems. Chi-squared tests revealed statistically significant sex differences for reading. Boys were more commonly found in the slow reading group at each age ( $p < 0.05$ ). There were also significantly more boys in the group with behaviour problems reported by parents at age seven and in the groups with behaviour problems reported by teachers at ages nine and 11.

##### Language delay and later mean intelligence

The means and standard deviations for the Verbal, Performance and Full-scale IQs among children for whom complete data were available at each age for each of the groups with comprehension delay, expressive delay and general language delay, and for the remainder of the sample, are set out in Table II.

The group-by-time effects were not statistically significant for any of the IQ measures. However, the between-group F

TABLE I  
Results from seven-, nine- and 11-year assessments

Measure	Age (yrs)	N	Mean	SD	Cut-off point	Below, at or above cut-off point			
						Boys N	Girls N	Total N	%
Verbal	7	950	105.7	15.13	≤ 89	68	57	125	13
	9	953	103.1	16.09		92	79	171	18
	11	919	104.4	15.72		74	82	156	17
Performance IQ	7	951	106.8	14.42	≤ 89	43	50	93	10
	9	953	105.2	15.20		63	57	120	13
	11	917	110.9	15.89		42	34	76	8
Full-scale IQ	7	950	106.9	14.52	≤ 89	57	47	104	11
	9	953	104.4	15.54		78	69	147	15
	11	917	108.2	15.76		56	52	108	12
Burt word-reading score	7	942	29.4	13.34	≤ 12	74*	26	100	11
	9	952	53.9	19.01	≤ 30	56*	15	71	8
	11	919	72.4	20.25	≤ 42	55*	22	77	8
Rutter Parent Scale	7	915	8.8	5.57	≥ 13	123*	85	208	23
	9	907	8.1	5.41		98	69	167	18
	11	895	7.2	5.48		71	56	127	14
Rutter Teacher Scale	7	940	4.0	4.81	≥ 9	79	55	134	14
	9	927	4.3	5.13		97*	42	139	15
	11	906	3.8	5.02		82*	43	125	14

\*Indicates significant sex difference ( $p < 0.05$ ).

TABLE II  
Language delay at age three and mean IQ at ages seven, nine and 11 years

IQ category	Age (yrs)	1. Comprehension delay (N = 20)		2. Expressive delay (N = 19)		3. General language delay (N = 19)		4. Remainder of sample (N = 812)		Between-group F ratios	Between-group comparisons 1:2 1:3 1:4 2:3 2:4 3:4					
		Mean	SD	Mean	SD	Mean	SD	Mean	SD							
Verbal	7	90.0	10.59	93.5	11.90	75.5	21.88	107.4	13.86	45.0	NS	*	*	*	*	*
	9	90.1	15.49	88.9	16.24	74.6	23.46	104.5	14.50							
	11	92.4	12.01	91.3	13.28	79.3	25.21	105.5	14.71							
Performance	7	99.4	18.77	95.5	12.86	88.0	18.12	108.0	13.72	25.7	NS	*	*	*	*	*
	9	92.3	16.00	94.3	13.49	84.3	21.11	106.6	14.02							
	11	101.2	18.80	98.8	15.73	91.5	22.20	111.8	15.06							
Full-scale	7	94.2	13.55	93.7	12.04	80.3	19.22	108.4	13.43	41.1	NS	*	*	*	*	*
	9	91.0	16.08	90.6	14.90	77.4	21.50	105.9	14.00							
	11	96.6	15.06	94.2	13.79	83.8	24.45	109.3	14.67							

\* $p < 0.05$ .

TABLE III  
Language delay at age three and mean Burt word-reading scores at ages seven, nine and 11 years

Age (yrs)	Comprehension delay (N=21)		Expressive delay (N=14)		General language delay (N=19)		Remainder of sample (N=806)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
7	17.2	7.62	17.1	7.99	12.8	10.82	30.7	12.89
9	39.4	14.04	35.4	12.26	32.2	16.73	55.2	18.36
11	54.4	15.65	54.2	18.00	46.6	23.32	73.7	19.29

F ratio (group by time) ~ 3.71.



TABLE IV  
Language delay at age three and mean Rutter Behaviour Scale scores at ages seven, nine and 11 years

Rutter Scale	Age (yrs)	1. Comprehension delay (N = 20)		2. Expressive delay (N = 16)		3. General language delay (N = 16)		4. Remainder of sample (N = 746)		Between-group F ratios	Between-group comparisons					
		Mean	SD	Mean	SD	Mean	SD	Mean	SD		1:2	1:3	1:4	2:3	2:4	3:4
Parent	7	11.2	6.04	8.8	3.94	15.1	6.70	8.7	5.45	9.80	NS	*	*	*	NS	*
	9	10.6	6.49	8.2	4.18	14.6	6.23	7.9	5.28							
	11	8.0	6.13	7.6	5.64	11.9	6.55	6.8	5.21							
Teacher	7	9.5	7.17	4.1	4.89	9.2	7.13	3.7	4.52	14.11	*	NS	*	*	NS	*
	9	6.7	6.68	5.3	4.53	7.9	7.23	4.0	4.96							
	11	7.2	8.13	5.4	5.11	7.1	5.90	3.7	4.89							

\* $p < 0.05$ .

ratios were significant, indicating that there were significant differences among the means on each IQ scale. *Post hoc* tests showed that all the language-delay groups had significantly lower mean IQs than the remainder of the sample. For each IQ type there were no significant differences between the groups with comprehension delay only and expressive delay only. However, those with general language delay gained a significantly lower mean IQ than the two other language-delay groups on all IQ scales.

#### *Language delay and later mean reading scores*

The means and standard deviations for the Burt Word Reading Test scores among children for whom complete data were available at each age for the three-year language-delay groups and the remainder of the sample are set out in Table III.

The group-by-time interaction effect for reading was statistically significant (approx.  $F(6,1722) = 3.71$ ). The reading scores for children in all the groups with delayed language increased at a significantly slower rate than those for the remainder of the sample over the four-year period from ages seven to 11. It is not appropriate to compare the mean scores of the language-delayed groups with each other because of the interaction effect that was found.

The mean reading scores of the groups delayed in either verbal comprehension or expression only at age 11 were similar to those of children two years younger. The

mean 11-year reading score of those with general language delay was similar to the mean of children more than 2½ years younger.

#### *Language delay and mean behaviour-problem scores*

The means and standard deviations for the Rutter Behavioural Scale scores among children for whom complete data were available for the language-delay groups and the remainder of the sample are set out in Table IV.

The group-by-time interaction effect for the behaviour scores was not statistically significant. However, there were statistically significant group effects for both measures. *Post hoc* tests on the Rutter Parent Scale showed that the groups with comprehension delay and general language delay, but not those with expressive delay, had significantly higher mean behaviour-problem scores than the remainder of the sample. The group with general language delay had a significantly higher mean behaviour score than the groups with comprehension delay and expressive delay. There was no significant difference in the scores for behaviour problems between those with delayed comprehension and expression.

*Post hoc* tests on the Rutter Teacher Scale again showed that the groups with comprehension and general delay, but not those with expressive delay, had significantly higher mean behaviour scores than the remainder of the sample. Those with comprehension delay had a significantly higher mean behaviour-

problem score than those with expression delay. Children with general delay had a significantly higher mean score than those with expressive delay. Children with general delay had a significantly higher mean score than those with expressive delay, but not than those with comprehension delay.

*Language delay, low IQ and reading delay*  
The numbers and proportions of children in each of the three-year language-delay groups and the remainder of the sample who had low IQs ( $\leq 89$ ) or low reading scores are set out in Table V.

All the over-all chi-squared tests for differences among the groups were significant. *Post hoc* tests showed that at ages seven and nine there were significantly more children from all three language-delay groups in each of the three categories of low IQ and in the reading-delay groups compared to the remainder of the sample. The same result was obtained from the group with general language delay at age 11. At age 11 the groups with comprehension delay and expressive delay had a significantly higher proportion of children in the low Verbal IQ group and the low Full-scale IQ group. There were no significant differences between the groups with either comprehension delay or expressive delay and the remainder of the sample in the proportion of children with a low Performance IQ or with delayed reading at age 11.

The proportions of children with low IQ or reading delay in the groups with delay in comprehension or expressive language were similar. A higher proportion of children with general language delay had a low IQ or a reading delay. About three times as many children with comprehension delay or expressive delay (59 to 67 per cent), and more than three times as many children with general language delay (71 to 86 per cent), had either a low IQ or a reading delay in comparison with the remainder of the sample (19 to 22 per cent).

#### *Language delay and later behaviour problems*

Over-all chi-squared tests were statistically significant for the Rutter

Teacher Scale at age seven, for both Teacher and Parent Scales at age nine and for the Parent Scale at age 11. *Post hoc* tests showed that children with general language delay were significantly different from the remainder of the sample on the Teacher Scale at age nine and on the Parent Scale at age 11. No other differences between the language-delay groups and the remainder of the sample were statistically significant (Table VI).

*Language delay and family disadvantage*  
The sample was divided according to whether the children scored 0 or 1, or 2 or more, on the family disadvantage index. The results for each language-delay group and the remainder of the sample are shown in Table VII.

The over-all chi-square was significant ( $p < 0.05$ ). *Post hoc* tests indicated that significantly more children with comprehension delay and general delay had a score of 2 or more on the family disadvantage index. The prevalence of family disadvantage in these groups was more than twice that for the remainder of the sample.

#### *Supplementary analyses*

Because there were significantly more boys than girls in the language-delay groups and because it was found that two of the three language-delay groups had significantly greater numbers of children with a score of 2 or more on the family disadvantage index, the results from the intelligence and reading scores (Tables II and III) were re-analysed, using analysis of covariance. The sex of the child and the score on the family disadvantage index were used as covariates. The effect of the covariates as dependent variables was significant in each case ( $p < 0.05$ ). The F ratios from the analysis of covariance for group effects were reduced, but all remained significant ( $p < 0.05$ ). This showed that the differences in mean IQ and reading scores were not due to sex differences or family disadvantage in the language-delay groups.

In order to test whether the association between language delay and increased behaviour problems was mediated by sex differences, family disadvantage or

TABLE V  
Language delay at age three, low IQ and reading delay at ages seven, nine and 11 years

	Comprehension delay N %	Expressive delay N %	General language delay N %	Remainder of sample N %	Over-all $\chi^2$ (3df)**
Age seven	(N=24)	(N=22)	(N=22)	(N=875)	
Low Verbal IQ	13 54*	12 54*	15 68*	83 9	137.9
Low Performance IQ	8 33*	6 27*	11 50*	68 8	66.7
Low Full-scale IQ	7 29*	10 45*	16 72*	70 8	129.2
Low reading score	9 38*	7 32*	14 64*	69 9	92.9
Any of the above†	16 67	13 59	19 86	178 20	
Age nine	(N=23)	(N=21)	(N=25)	(N=877)	
Low Verbal IQ	12 52*	13 62*	18 72*	125 14	104.4
Low Performance IQ	10 43*	7 33*	15 60*	86 10	86.1
Low Full-scale IQ	12 52*	11 52*	21 84*	101 12	146.9
Low reading score	5 22*	6 28*	10 40*	49 6	65.9
Any of the above†	15 65	14 67	21 84	19 22	
Age 11	(N=22)	(N=19)	(N=21)	(N=848)	
Low Verbal IQ	12 55*	9 47*	14 67*	118 14	75.8
Low Performance IQ	4 18	4 21	9 43*	57 7	63.4
Low-Full-scale IQ	8 36*	7 36*	13 62*	78 9	81.3
Low reading score	4 18	4 21	11 54*	57 7	62.7
Any of the above†	13 59	11 58	15 71	162 19	

\*Indicates difference between these groups and remainder of sample significant ( $p < 0.05$ ).

\*\*All over-all  $\chi^2$  tests significant ( $p < 0.001$ ).

†No statistical tests were carried out on these data.

TABLE VI  
Language delay at age three and behaviour problems at ages seven, nine and 11 years

	Comprehension delay N %	Expressive delay N %	General language delay N %	Remainder of sample N %	Over-all $\chi^2$ (3df)
Age seven	(N=24)	(N=22)	(N=25)	(N=873)	
Parent scale	8 33	3 14	9 36	186 21	6.94
Teacher scale	11 46	2 9	8 32	112 13	29.13**
Either scale†	16 67	4 18	13 52	257 29	
Age nine	(N=23)	(N=21)	(N=25)	(N=872)	
Parent scale	16 26	5 24	18 72	142 16	9.06**
Teacher scale	7 30	6 28	8 32*	114 13	14.91**
Either scale†	11 48	9 43	14 56	223 26	
Age 11	(N=23)	(N=19)	(N=23)	(N=848)	
Parent scale	3 13	2 11	9 39*	112 13	14.53**
Teacher scale	6 26	4 21	5 22	110 13	6.22
Either scale†	6 26	4 21	12 52	197 23	

\*Indicates difference between these groups and remainder of sample significant ( $p < 0.05$ ).

\*\*Indicates over-all  $\chi^2$  significant ( $p < 0.05$ ).

†No statistical tests were carried out on these data.

TABLE VII  
Language delay at age three and Disadvantage Index scores

Disadvantage Index score	Comprehension delay (N=27) N %	Expressive delay (N=24) N %	General language delay (N=27) N %	Remainder of sample (N=949) N %
0 or 1	11 41	16 67	10 37	685 72
2 or more	16 59*	8 33	17 63*	264 28

\* $p < 0.05$ .

reading delay, the analyses of the behaviour-problems data (Table IV) were repeated, using analysis of covariance. In these analyses the sex of the child, the score on the family disadvantage index and the child's reading score were used as covariates. The effects of the covariates were significant in each case ( $p < 0.05$ ). The analysis resulted in a reduced F ratio for group effects, but they remained significant ( $p < 0.05$ ). This showed that the increased scores for behaviour problems in the groups with language delay were not due to sex differences, family disadvantage or poor reading ability.

### Discussion

This study was an extension of an earlier one (Silva *et al.* 1983) which described the development of language-delayed children to age seven. It is the first large-scale study of children with language delay in the preschool years to have followed them to age 11. However, Aram and co-workers (1984) recently reported a small-scale study which followed 20 of an original sample of 47 language-delayed children 10 years after a preschool assessment. A high proportion of those children were found to be experiencing educational difficulties.

Although the three follow-up studies reviewed by Stevenson (1984) were in some ways comparable, there was an important difference between them. This related to the method of identifying children with language delay and the resulting proportion of children so defined. The Newcastle study (Fundudis *et al.* 1979) and the Waltham Forest study (Richman *et al.* 1983) defined 3 per cent of their sample as having language delay. The Dunedin study used a different definition which identified approximately 2.6 per cent as delayed in comprehension, 2.3 per cent as delayed in expressive language and 2.6 per cent as having a general language delay (a total of 7.5 per cent). Despite the larger proportion of children defined as language-delayed in that study, all three groups were found to be disadvantaged in later development in comparison with the remainder of the sample.

Analysis of the mean IQ results showed that each group with language delay had

significantly lower mean IQs at each age than the remainder of the sample. As would be expected (Stevenson 1984), the Verbal IQs were lower than the Performance IQs. Those with general language delay had the lowest mean IQs (Full-scale 77 to 83); at each age these were significantly lower than those of the two other language-delay groups. While it was not appropriate to compare the mean reading scores of all the language-delay groups, the analysis showed that all three groups with language delay increased their reading scores at a significantly slower rate than the remainder of the sample.

In order to describe further the long-term development of children with early language delay, the prevalence of later low IQ and low reading scores were investigated (see Table V). At ages seven and nine, all three language-delay groups were significantly over-represented in each type of low-IQ group, and in the group with low reading scores, compared to the remainder of the sample. At age 11 these results still held for the group with general language delay. At age 11 both groups with delayed comprehension or expression had significantly higher proportions of children with low Verbal IQs and low Full-scale IQs. However, they were not significantly over-represented in the groups with either low Performance IQs or with low-reading scores. Between 58 and 67 per cent of those with either comprehension or expression delay had a low IQ and/or a low reading score at ages seven, nine and 11, in contrast to 19 to 22 per cent of the remainder of the sample. Between 71 and 86 per cent of children with general language delay had a low IQ and/or reading difficulties across the three ages studied. These results further describe the long-term significance of early language delay and support the conclusions from the earlier studies reviewed by Stevenson (1984), that children with early language delay are at high risk of later low intelligence and poor reading progress. The results also suggest that this applies not only at ages seven and eight, but also at ages nine and 11. Thus language delay appears to be predictive of longer-term low intelligence and poor reading progress.

The results in respect of behaviour problems were less clear-cut. Analysis of the total scores on the Rutter Scales showed that children with comprehension delay and general language delay, but not those with expressive language delay, had significantly higher scores for behaviour problems on both the Parent and Teacher Scales over the four-year period from seven to 11. Thus only two of the three language-delay categories were predictive of behaviour problems from ages nine to 11. When behaviour problems were analysed in terms of the proportion defined as having a later behaviour problem, only the group with general language delay had significantly more children with problems, and then only on one scale (Teacher) at age nine and one scale (Parent) at age 11. Over-all, the results suggest that general language delay and comprehension delay (but not expressive language delay) is predictive of higher mean scores on both of the Rutter Behaviour Scales at ages seven, nine and 11.

Of the three groups with language delay, that with general language delay was consistently the most disadvantaged in terms of the follow-up measures, followed by the group with comprehension delay. Compared to the other two groups, those with general language delay had significantly lower mean IQs over the three ages. They also had significantly higher mean scores on the Rutter Parent Scale than either of the other two groups over the three ages. They also had significantly higher scores on the Teacher Scale than the children with expressive language delay over the three ages. This tends to support the contention of Stevenson (1984) that general language delay is the type most predictive of later problems.

This study also showed that significantly more boys than girls had language delay, and that the two most developmentally delayed groups (general language and comprehension) had significantly higher scores on the disadvantage index. Because of this, the intelligence and reading-test results were re-analysed, using analysis of covariance, with the child's sex and the disadvantage score as covariates. This produced a

reduction in the size of the between-group effect, but the *F* ratios still remained significant. This indicated that the earlier results held, even when the possible confounding effects of the two background variables were controlled. The results also held for behaviour problems when the above variables and the child's reading score were included in an analysis of covariance. This suggested that the higher mean scores for behaviour problems in the groups with comprehension delay and general language delay were not due to sex differences, family disadvantage or poor reading ability, although these were contributing factors.

The main conclusion from this study is that children with early language delay are an exceptionally high-risk group, since they turned out to have lower intelligence, lower reading scores and higher scores for behaviour problems, not only to ages seven and eight, but also to 11. Stevenson (1984), extrapolating from studies from the British National Child Development Study, speculated that language screening at three years 'is likely to identify children who will be at risk of experiencing a wide range of difficulties, at least until they leave school' (p. 535). The Dunedin results give support to that view to age 11. The Dunedin sample is now being assessed at age 13 and there are plans for further assessment at ages 15 and 17. These further studies should clarify the long-term implications of early language delay.

Some authors (e.g. Stark *et al.* 1983) have argued that ascertainment of children with language delay should be at the age of seven to eight because of the instability of early language delay. The results from the Dunedin study do not support this view as they show very high rates of later developmental disadvantage and behavioural problems among children identified as having a language delay as early as age three.

While it is possible to identify children with language delay in the preschool years, the effectiveness of intervention programmes and family support is less clear. As pointed out by Stevenson (1984), there is no shortage of programmes and there is evidence to support the effectiveness of some of them (e.g.



Cooper *et al.* 1979), but further research is needed. The results from the Dunedin study suggest that if such effective programmes are not used, then the majority of children with early language delays will very likely experience long-lasting problems in cognitive, educational and behavioural development in comparison with those whose early language development was not delayed.

Accepted for publication 13th November 1986.

#### Acknowledgements

The Dunedin Multidisciplinary Health and Development Study is supported by the Medical

Research Council of New Zealand and the Departments of Education and Health, and involves several departments of the University of Otago. Many of the data are gathered by voluntary workers from the Dunedin community. The authors are indebted to the many people whose contributions make this continuing study possible.

#### Authors' Appointments

\*Phil A. Silva, Director;

Sheila Williams;

Rob McGee;

The Dunedin Multidisciplinary Health and Development Research Unit, Department of Paediatrics and Child Health, University of Otago Medical School, P.O. Box 913, Dunedin, New Zealand.

\*Correspondence to first author.

#### SUMMARY

A large sample of Dunedin (New Zealand) children were assessed at age three to identify those with language delay. 2.6 per cent were defined as delayed in verbal comprehension only, 2.3 per cent as delayed in verbal expression only, and 2.3 per cent as delayed in both ('general language delay'). Most of these children, and the remainder of the sample, were assessed for intelligence, reading and behaviour problems at ages seven, nine and 11. Those in every language-delay group had significantly lower mean IQs and lower mean reading scores than the remainder of the sample. They also more often had a low IQ or a lower reading score at ages seven and nine and a lower Verbal and Full-scale IQ at 11. The groups with delayed verbal comprehension and general language delay had significantly more behaviour problems than the remainder of the sample. The group with general language delay was consistently the most disadvantaged in later intelligence, reading and behaviour. Two of the language-delay groups (comprehension and general language) had significantly higher scores on a family disadvantage index. The results of this study confirm the importance of early language delay as a predictor of lower than average intelligence and reading ability and increased behaviour problems.

#### RÉSUMÉ

*Etude longitudinale des enfants avec retard de langage à trois ans: problèmes ultérieurs d'intelligence, de lecture et de comportement*

Un échantillon important des enfants de Dunedin (Nouvelle Zélande) a été examiné à l'âge de trois ans, à la recherche d'un retard de langage. 2.6 pour cent ont été définis comme en retard sur le seul plan de la compréhension verbale, 2.3 pour cent sur le seul plan de l'expression verbale et 2.3 pour cent sur les deux plans (retard global de langage). La plupart de ces enfants et du reste de l'échantillon ont été évalués pour l'intelligence, la lecture et les problèmes de comportement aux âges de sept, neuf et onze ans. Dans chaque groupe de retard de langage, on notait une moyenne de Q.I. et une moyenne d'efficacité de lecture plus basses que dans le reste de l'échantillon. La fréquence des Q.I. et des scores de lecture bas aux âges de sept et neuf ans, les échelles verbales et totales basses à onze ans étaient également plus élevées. Les groupes avec retard de compréhension verbale et retard global de langage présentaient plus de problèmes de comportement que le reste de l'échantillon. Le groupe avec retard global de langage était constamment le plus désavantagé en retard intellectuel, lecture et comportement. Deux des groupes retardés (retard en compréhension et retard global de langage) avaient un score plus élevé à un index de désavantage familial. Les résultats de cette étude confirment l'importance des retards précoces de langage comme prédicteurs d'une intelligence et d'une intelligence et d'une aptitude de lecture au dessous de la moyenne, et de problèmes comportementaux accrus.

#### ZUSAMMENFASSUNG

*Eine Langzeitstudie bei Kindern mit entwicklungsbedingter Sprachverzögerung im Alter von drei Jahren: Spätere Intelligenz-, Lese- und Verhaltensprobleme*

In Dunedin (Neu Seeland) wurde eine große Gruppe von Kindern im Alter von drei Jahren untersucht, um diejenigen mit einer Sprachverzögerung herauszufinden. 2.6 Prozent hatten nur eine Verzögerung im verbalen Verständnis, 2.3 Prozent nur im verbalen Ausdruck und 2.3 Prozent in beidem (allgemeine Sprachverzögerung). Die meisten dieser Kinder und die übrigen der Gruppe wurden im Alter von sieben, neun und 11 Jahren mit einem Intelligenztest und auf Lese- und Verhaltensprobleme untersucht. Die Kinder aus den Sprachverzögerungsgruppen hatten signifikant niedrigere mittlere IQs und Lesescores als die übrigen Kinder. Mit sieben und neun Jahren hatten sie häufiger einen niedrigen IQ oder einen niedrigeren Lesescore und mit 11 Jahren einen niedrigeren verbalen und allgemeinen IQ. Die Kinder mit einer Verzögerung im verbalen Verständnis und mit einer allgemeinen Sprachverzögerung hatten signifikant häufiger Verhaltensprobleme als die übrigen. Die Kinder mit einer allgemeinen Sprachverzögerung waren immer im Verlauf ihrer Entwicklung in

ihrer Intelligenz, beim Lesen und im Verhalten am stärksten beeinträchtigt. Zwei der Sprachverzögerungsgruppen (Verständnis und allgemeine Sprache) hatten signifikant höhere Scores bei einem Index für familiäre Benachteiligung. Die Ergebnisse dieser Studie unterstreichen die Bedeutung der früherkannten Sprachverzögerung als Hinweis auf eine unterdurchschnittliche Intelligenz und Lesefähigkeit und auf erhöhte Verhaltensprobleme.

## RESUMEN

*Estudio longitudinal de niños con retraso en el desarrollo de lenguaje a los tres años de edad: inteligencia, lectura y problemas de desarrollo posterior*

Una gran muestra de niños en Dunedin (Nueva Zelanda) fue evaluado a la edad de tres años para identificar los que tenían un lenguaje retrasado. 2·6 por ciento fueron definidos como retrasados sólo en la comprensión verbal, 2·3 en la expresión y 2·3 en ambas ('retraso general del lenguaje'). La mayoría de los niños y el resto de la muestra fueron evaluados en su inteligencia y problemas de lectura y de comportamiento a la edad de siete, nueve y 11 años. Los que tenían un retraso en el lenguaje de cualquier tipo tenían un promedio de CI significativamente más bajo y un puntaje promedio de lenguaje más bajo que el resto de la muestra. También con mayor frecuencia tenían un CI bajo y un puntaje bajo en la lectura a las edades de siete y nueve años y un CI más bajo global y del lenguaje a la edad de 11 años. Los grupos con un retraso en la comprensión verbal y en el lenguaje en general tenían significativamente mayores problemas de desarrollo que el resto de la muestra. El grupo con retraso general del lenguaje tenía consistentemente mayor desventaja en la inteligencia posterior, la lectura y el comportamiento. Dos de los grupos con retraso en el lenguaje (comprensión y global) tenían un puntaje significativamente más alto para un índice familia desventajoso. Los resultados de este estudio confirman la importancia de un retraso precoz en el lenguaje como predictor de un puntaje más bajo de inteligencia y de capacidad de lectura y un aumento de problemas de conducta.

## References

- Aram, D. M., Ekelman, B. L., Nation, J. E. (1984) 'Preschoolers with language disorders 10 years later.' *Journal of Speech and Hearing Research*, 27, 232-244.
- Cooper, J., Moodley, M., Reynell, J. (1979) 'The developmental language programme. Results from a five year study.' *British Journal of Disorders of Communication*, 14, 57-59.
- Elley, W. B., Irving, J. C. (1972) 'A socio-economic index for New Zealand based on levels of education and income from the 1966 census.' *New Zealand Journal of Educational Studies*, 7, 155-167.
- Everitt, B. S. (1977) *The Analysis of Contingency Tables*. London: Chapman & Hall.
- Fundudis, T., Kolvin, I., Garside, R. F. (Eds.) (1979) *Speech Retarded and Deaf Children: Their Psychological Development*. London: Academic Press.
- Gilmore, A., Croft, C., Reid, N. (1981) *Burt Word Reading Tests, New Zealand Revision: Teachers Manual*. Wellington: N.Z. Council for Educational Research.
- McGee, R. O. (1984) *Response Rates at Phase IX of the Dunedin Multidisciplinary Health and Development Research Unit: Unpublished Report from the Dunedin Multidisciplinary Health and Development Research Unit*. Available from the University of Otago Medical Library.
- Silva, P. A. (1982) *A Thousand New Zealand Children. Their Health and Development from Birth to Seven*. Medical Research Council of New Zealand, Special Report Series No. 8.
- Williams, S. M. (1983) 'Parents' and teachers' perception of behaviour problems in seven year old children.' *Exceptional Child*, 30, 151-161.
- — (1984) 'Behaviour problems in a population of seven year old children: prevalence, stability and types of disorder.' *Journal of Child Psychology and Psychiatry*, 25, 251-259.
- Williams, S., Bradshaw, J., Chapel, J. L., Robins, A., Silva, P. A. (1985) 'The Rutter scale for completion by teachers: factor structure and relationship with cognitive abilities and family adversity for a sample of New Zealand children.' *Journal of Child Psychology and Psychiatry*, 26, 727-739.
- Reynell, J. (1969) *Reynell Developmental Language Scales*. Windsor: N.F.E.R.
- Richman, N., Stevenson, J., Graham, P. (1983) 'The relationship between language, development, and behaviour.' In Schmidt, M. H., Renschmidt, H. (Eds.) *Epidemiological Approaches in Child Psychiatry II*. New York: Thieme-Stratton.
- Rutter, M., Tizard, J., Whitmore, K. (1970) *Education, Health and Behaviour*. London: Longman.
- Scottish Council for Research in Education (1976) *The Burt Word Reading Test, 1974 Revision*. London: Hodder & Stoughton.
- Silva, P. A. (1980) 'A study of the prevalence, stability and significance of developmental language delays in preschool children.' *Developmental Medicine and Child Neurology*, 22, 768-777.
- (1981) 'A simple prose reading test to identify seven year olds with reading problems: a report from The Dunedin Multidisciplinary Child Development Study.' *National Education*, 63, 36-38.
- McGee, R. O., Williams, S. M. (1983) 'Developmental language delay from three to seven years and its significance for low intelligence and reading difficulties at age seven.' *Developmental Medicine and Child Neurology*, 25, 783-793.
- SPSS-X (1983) *SPSS-X Users Guide*. New York: McGraw-Hill.
- Stark, R. E., Mellitis, E. D., Tallal, P. (1983) 'Behavioural attributes of speech and language disorders.' In Ludlow, C. L., Cooper, J. A. (Eds.) *Genetic Aspects of Speech and Language Disorders*. New York: Academic Press.
- Stevenson, J. (1984) 'Predictive value of speech and language screening.' *Developmental Medicine and Child Neurology*, 26, 528-538.
- Wechsler, D. (1974) *The Wechsler Intelligence Scale for Children—Revised*. New York: Psychological Corporation.