JSLHR, Volume 40, 556–566, June 1997

Late Talkers at 2: Outcome at Age 3

Leslie Rescorla
Bryn Mawr College
Bryn Mawr, PA

Julie Roberts
University of Vermont
Burlington

Katherine Dahlsgaard
University of Pennsylvania
Philadelphia

Age 3 follow-up data are presented for a sample of 34 toddlers diagnosed between the ages of 24 and 31 months with expressive type specific language impairment (SLI-E). At age 3, the late talkers scored significantly lower on all language measures than 21 comparison peers matched at intake on age, SES, and nonverbal ability. When seen at follow-up, the former late talkers scored in the average range on the Expressive One-Word Picture Vocabulary Test (EOWPVT) and on the Reynell Expressive Language Scale, but more than 1.5 *SD*s below age expectations in MLU and on Scarborough's (1990a) IPSyn. The proportion of late talkers performing in the average range at follow-up varied markedly as a function of measure used (EOWPVT: 79%, Reynell: 58%, MLU: 35%, and IPSyn: 24%), indicating that the late talkers made more rapid progress in lexical development and in the use of language to define, explain, and describe than they did in the areas of syntactic and morphological development. The only significant predictor of age 3 outcome was intake expressive language level, with toddlers who had been more severely delayed in expressive language at intake relative to age level having the worst outcomes at age 3.

KEY WORDS: late talkers, SLI-E, language outcome

Young children with delayed expressive language are known to be at risk for continuing language problems, emotional/behavioral disorder, and later learning disabilities (Aram, Ekelman, & Nation, 1984; Bishop & Adams, 1990; Bishop & Edmundson, 1987; Catts, 1991; Richman, Stevenson, & Graham, 1982; Scarborough, 1990b; 1991; Silva, Williams, & McGee, 1987; Tallal, 1988; Tomblin, Freese, & Records, 1992). However, not all young children who are slow to talk develop into youngsters with language impairment or learning disabilities. Many late talkers catch up in language during the preschool years and are indistinguishable from their peers in later language, academic achievement, and emotional/behavioral adjustment (Bishop & Adams, 1990; Paul, Laszio, McFarland, & Midford, 1993; Rescorla, 1993; Whitehurst, Smith, Fischel, Arnold, & Lonigan, 1991; Whitehurst, Fischel, Arnold, & Lonigan, 1992; Whitehurst & Fischel, 1994).

Most follow-up studies of children with language delays have involved youngsters first identified at age 3 or 4 (Bishop & Edmundson, 1987; Fundudis, Kolvin, & Garside, 1979; Klackenberg, 1980; Richman, Stevenson, & Graham, 1982; Silva et al., 1987; Tallal, 1988). However, five cohorts of late-talking toddlers have been the subject of follow-up study in the past decade. These studies have focused on late talkers identified at around age 2, the age when parents generally begin to become concerned if their child is not talking. As will emerge from the review presented below, results from each of these research programs

 1092-4388/97/4003-0556

have added to our knowledge of what happens to late-talking toddlers as they get older, but there are still many aspects of this basic question for which we do not currently have answers.

The San Diego Sample

The San Diego sample (Thal & Tobias, 1992; Thal, Tobias, & Morrison, 1991) consisted of ten 18- to 32-month-old toddlers who were in the lowest 10% of expressive vocabulary development by maternal report on a precursor of the MacArthur Communicative Development Inventory (Fenson et al., 1993) (e.g., fewer than 65 words and few two-word combinations; Thal & Bates, 1988). The children were described as normal in nonverbal cognitive ability, although no standardized test scores were reported. Although, as a group, the late talkers were comparable to age-matched controls in receptive language as assessed by parent report and by a two-choice picture identification task, 40% of the sample had significant comprehension delays (Thal et al., 1991).

At one-year follow-up, six children ("late bloomers") appeared to have caught up in expressive language and four were still delayed ("truly delayed"), based on parent report of vocabulary and MLU from a speech sample. Results indicated (Thal et al., 1991; Thal & Tobias, 1992) that the late bloomers could be discriminated from the truly delayed children by their age-appropriate receptive language and communicative gesture production at intake. Vocabulary size, MLU, and word-to-vocalization use at intake did not predict outcome (Thal & Tobias, 1992).

The Portland Sample

The most complete report on language outcome for Paul's Portland sample (Paul, 1993) summarizes data for 37 children. These late talkers were identified when they were between 20 and 34 months old because they had fewer than 50 words by parent report using Rescorla's (1989) Language Development Survey. All late talkers had Bayley Mental Development Index scores above 85 and were equivalent to 32 age-matched control children on nonverbal items from the Bayley (Paul, 1991). Roughly one-quarter of the late talkers were more than 6 months delayed in receptive language according to maternal report on the Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984; Paul, 1991).

Follow-up at age 3 (Paul, 1993) indicated that the late talkers did not differ from the comparison group in receptive language or in expressive vocabulary. However, half of the late talkers scored below the 10th percentile on the Goldman-Fristoe Test of Articulation, 60% were below the 10th percentile in Developmental Sentence Score (DSS; Lee, 1974), and 37% were delayed in both areas. When followed up at age 4 (Paul, 1993), 34% of the late talkers were below the 10th percentile in TOLD-P Word Articulation, 47% were below the 10th percentile in DSS, and 16% were still delayed in both areas. Furthermore, the children in the late talker group who had caught up to the comparison group in MLU were significantly less advanced than their normal peers in grammatical morphology (Paul & Alforde, 1993).

Preliminary prediction results from regression analyses indicated that the older the child at time of diagnosis (within the age range of 20 to 34 months), the less positive the outcome (Paul, 1993). In addition, although boys outnumbered girls in the late talker group by a large margin (75% to 25%), girls had a lower probability of moving into the normal range in syntax than boys. Paul (1993) did not indicate whether or not initial levels of receptive language predicted language outcome.

The Wisconsin Sample

Outcome for four late-talking toddlers was recently reported by Ellis Weismer, Murray-Branch, and Miller (1994). The late talkers had normal nonverbal ability and approximately normal receptive language. At 25 to 26 months, they had expressive vocabulary below the 10th percentile on the MacArthur Communicative Development Inventory (Fenson et al., 1993) (range of 25 to 87 reported words), but all were producing some word combinations, had phonetic inventories within the normal range, and had MLUs ranging from 1.03 to 1.22 from a naturalistic speech sample. At the 31- to 33-month follow-up, three of the four toddlers had reported vocabularies in the normal range (more than 500 words), but the fourth had a reported vocabulary of only 120 words. By 34 to 35 months, one of the four children had a normal MLU, two others had MLUs more than 1 *SD* below average, and the fourth child (the boy with the smallest vocabulary) was more than 2 *SD*s below average in MLU.

When Ellis Weismer and colleagues (Ellis Weismer et al., 1994) examined predictors of outcome, they found that the child with the best outcome had the worst receptive language at intake, whereas the child with the poorest outcome had manifested the largest vocabulary and the best receptive language at 25 to 26 months and had received 3 months of intervention. The authors concluded that outcome in this multiple case study could not be predicted by intake vocabulary size, MLU, receptive language, or symbolic play skills.

The New York Sample

The New York sample of late-talking toddlers (Fischel, Whitehurst, Caulfied, & DeBaryshe, 1989; Whitehurst et al., 1992; Whitehurst & Fischel, 1994) consisted of 22 children (age range: 24 to 38 months) with specific expressive language delay. All the children had IQs above 85 on the Leiter International Performance Scale (a nonverbal test), a Revised Peabody Picture Vocabulary Test score of at least 85 (within 1 *SD* of average), and performance on the Expressive One-Word Picture Vocabulary Test of no greater than 65 (2.33 *SD*s below average). Mean expressive vocabulary by parent report on a checklist was 17.54 words.

When followed up 5 months later, the late talkers were still about 10 months behind on Expressive One-Word Picture Vocabulary Test score, with only 35% of the group having scores in the average range (>85). The sample as a whole was within age expectations on the Verbal Expression subtest of the ITPA at follow-up, but 26% had scores below 85 (Whitehurst & Fischel, 1994). Outcome was predicted by reported vocabulary size at intake (Fischel et al., 1989), as well as by proportion of consonants to vowels in vocalizations during intake (Whitehurst et al., 1991).

By age 3:6 (years:months), 88% of the New York sample were in the average range (above 85) on the Expressive One-Word Picture Vocabulary score, and 96% were in this range by 5:6 (Whitehurst, Fischel, Arnold, & Lonigan, 1992). On the Verbal Expression subtest of the ITPA, 16% of the sample were 1 *SD* below average at 3:6, and only 7% were below average at age 5:6. Unfortunately, no naturalistic speech samples were analyzed in the New York study; thus no data on MLU or grammatical development are available for this sample of late-talking children with normal receptive language skills.

The Pennsylvania Sample

The last group of late-talking toddlers that has been discussed in the literature is Rescorla's Pennsylvania sample (Rescorla, 1993; Rescorla & Schwartz, 1990). Rescorla and Schwartz (1990) reported on the first 25 children from this sample for whom follow-up data were available. All subjects were diagnosed between 24 and 31 months of age with specific expressive language delay. The late talkers all had Bayley MDI scores of at least 85, Reynell Receptive Language Scale scores within 4 months of chronological age, and Reynell Expressive Language Scale score at least 6 months below age level. They also had fewer than 50 words or no word combinations by maternal report on the Language Development Survey (Rescorla, 1989).

In Rescorla and Schwartz (1990), data were reported for each child at a minimum of 8 months after intake. Sixteen children were 3 years old at follow-up, and the other 9 were 42 to 48 months old. Based on 100-utterance speech samples collected during free play with their mothers, the children had a mean MLU that was 1.68 *SD*s below age expectations in terms of benchmark scores from Scarborough (1990a). Only 10 of the 25 late talkers had MLUs within 1 *SD* of age expectations at follow-up. When their speech samples were analyzed using Scarborough's (1990a) IPSyn, their mean IPSyn z score was –3.06. Only 7 of the 25 were within 1 *SD* of age expectations on this measure of noun and verb phrase development, question/negation formation, and maturity of sentence structure. The only significant predictors of outcome reported by Rescorla and Schwartz (1990) were age and degree of language lag at intake. In general, late talkers who lagged farthest behind age expectations (these also tended to be the oldest toddlers) had the poorest outcomes (e.g., correlations of –.58 and –.70 with MLU and IPSyn respectively at follow-up).

Additional follow-up data from the Pennsylvania sample were presented by Rescorla (1993) and Rescorla and Dahlsgaard (1995). Data from standardized language tests indicated that the late talkers as a group were functioning in the average range on tests of syntax by age 5, but that the children in the SLI-E group continued to perform below the level of comparison group youngsters on most vocabulary, syntax, and formulation tests administered through age 8.

Summary

In summary, this review of outcome data from five cohorts of late talking toddlers strongly suggests that toddlers who are slow to talk are at risk for continuing expressive language delay to age 3. Predictors of outcome have not been consistent across previous studies. In one of the five projects reviewed (Ellis Weismer et al., 1994), no reliable predictors of outcome were found. In the other four studies, some significant predictors were identified: namely receptive language and gesture production (Thal & Tobias, 1992), age at intake (Paul, 1993; Rescorla & Schwartz; 1990), gender (Paul, 1993), expressive language level (Fischel et al., 1989; Rescorla & Schwartz, 1990), and consonantal repertoire (Whitehurst et al., 1991).

Of these five toddler cohorts, only the New York and Pennsylvania samples report outcomes for a significant number of late talkers with normal receptive language. In these two studies, fewer than half of the children with expressive type specific language delay had age-appropriate language by age 3. However, because speech sample data were not analyzed in the New York sample, the outcome of their subjects in terms of syntax and morphology is difficult to assess. Although Rescorla and Schwartz (1990) presented the most complete report to date on the outcome of late talkers with specific expressive language delay (SLI-E), their children ranged in age from 3 to 4 years at outcome, and no comparison group data were reported.

In the current report, we present follow-up data at age 3 for 34 subjects from the Pennsylvania sample of toddlers manifesting specific expressive language delay

(SLI-E) between 24 and 31 months. Their data at follow-up are measured against similar data collected from an age-matched comparison sample of 21 children who had been fully equivalent to the late talkers in age, SES, and Bayley nonverbal cognitive ability at intake. We address three principal questions in this report. First, do children who have a history of specific expressive language delay at 24 to 31 months differ from comparison children with normal language histories on standardized language tests and naturalistic speech indices at age 3? Second, what percentage of late talkers identified between 24 and 31 months with specific expressive language delay performed within the average range at age 3 on various lexical and syntactic measures? Finally, could outcome at age 3 be predicted from information collected on the late talkers at the 24-to-31-month intake?

Method

Subjects

Subjects for the study included 34 late talkers (33 boys and 1 girl) who were 24 to 31 months old at the time of intake and an age-matched comparison group of 21 toddlers (20 boys and 1 girl) with normal language. The late talkers whose age 3 outcomes are reported here consist of all subjects in Rescorla's Pennsylvania sample of 40 children for whom a 36-month naturalistic speech sample was collected during mother-child play with the Fisher Price Village. The six excluded youngsters were among the first subjects recruited into the sample, and they had either played with different toys ($n = 5$) or had defective tapes ($n = 1$). Of the 34 children in the present sample, 21 had been in the Rescorla and Schwartz (1990) sample of 3- to 4-year-olds. The comparison children described here are the 21 youngsters with normal language development in the Pennsylvania sample seen both at intake and at age 3.

Late talkers were recruited through newspaper advertisements, notices to pediatricians, and a local infant lab. All but one of the children in the study came from a two-parent, middle to upper-middle class white family (the mother of one late talker was divorced). All the late talkers had normal nonverbal abilities and age-adequate receptive language but significant delays in expressive speech. They had to have an MDI score of greater than 85 on the Bayley Mental Development Scale (Bayley, 1969). Additional selection criteria were a score within 3 months of chronological age on the Reynell Receptive Language Scale (Reynell, 1977) and a score at least 6 months below CA on the Reynell Expressive Language Scale. Comparison children met the same Bayley and Reynell Receptive Language Scale criteria and also had to have a Reynell Expressive Language Scale score within 3 months of CA.

All subjects in both groups met these group selection criteria except for three late talkers whose receptive language skills were 4 rather than 3 months delayed and one comparison child who had an expressive language age of 4 months below CA. Because none of these deviations was sufficient to place the subject into the contrasting group and the results of the study were unchanged when these children were excluded, they were included in their respective groups for all analyses.

Demographic information and test scores for the two subject groups appear in Table 1. As can be seen, the late talkers and the comparison children were essentially identical in age and Hollingshead SES score (maximum of 66). The two groups were also fully comparable in their total scores on the 19 nonverbal Bayley items above the basal level for all children (item #123). These items tap such skills as stacking blocks, putting puzzles together, drawing, and inserting pegs.

As the data in Table 1 show, the groups were significantly different in receptive language as measured by the Reynell Receptive Language Scale z score, $t(39.84) = -4.11$, $p < .001$. Although the late talkers had fully normal receptive skills for their age (26.85 month age level) (e.g., all were within .7 *SD*s of age expectations), the comparison children were advanced in receptive language (29.48 month age level). Of course, there was a striking difference in Reynell Expressive Language Scale z-score between the two groups, $t(40.07) = -14.53$, $p < .001$. On average, the late talkers were at the 16.82 months age level in expressive language (a lag of more than 9 months), whereas the comparison children were slightly advanced (26.48 months age level). All late talkers were at least 1.2 *SD*s below age expectations, and 26 of the 34 were 1.5 *SD*s or more below age level on the Reynell Expressive Language Scale.

As can be seen in Table 1, the late talkers had a mean reported vocabulary on Rescorla's (1989) Language Development Survey (LDS) of 20 words, in contrast to a mean reported vocabulary of 226 words for the comparison children, $t(22.68) = -13.56$, $p < .001$. All late talkers met

Table 1. Intake measures for late talkers and comparison children.

Measure	Late talkers		Comparison children	
	M	*SD*	*M*	*SD*
Intake age (months)	26.26	2.44	25.43	1.96
Hollingshead total	52.59	13.19	52.62	9.92
Bayley nonverbal items	13.62	3.13	13.05	3.84
Reynell receptive z-score	0.16	0.58	0.86	0.63***
Reynell expressive z-score	−1.72	0.48	0.30	0.52***
LDS vocabulary[a]	20.12	22.04	225.67	67.21***

[a]LDS = Language Development Survey (Rescorla, 1989)
***$p < .001$

Rescorla's (1989) criterion of fewer than 50 words or no word combinations on the LDS (one late talker had more than 50 words and five had a few word combinations).

Procedure

All subjects were seen for follow-up at age 3 in the company of their mothers. All of the comparison children and 24 of the late talkers were given the Leiter International Performance Scale, a nonverbal intelligence measure that taps color, form, number, and categorical concepts, and the Test of Auditory Comprehension of Language–Revised (TACL–R; Carrow-Woolfolk, 1985), a receptive language measure involving pictorial stimuli. The Expressive One-Word Picture Vocabulary Test (EOWPVT; Gardner, 1981), a single word-naming test employing pictures, was given to 28 of the late talkers and all the comparison children. The late talkers who did not receive these three outcome measures were the earliest subjects recruited into the project, and they had their age 3 visit before this test battery was finalized. The Reynell Expressive Language Scale was administered to 33 of the 34 late talkers (one late talker did not receive the Reynell because of administration error). At age 3, this scale involves picture labeling, defining words, describing pictures, and a global rating of the child's syntactic maturity. Finally, each child played with his or her mother for approximately 30 minutes using the Fisher Price Village, a toy that contains a wide variety of environments and equipment conducive to pretend play (e.g., a fire engine, a garage, a barber shop, a post office, and many toy figures, vehicles, and other assorted props). This play session was both videotaped and audiotaped. In addition, a speech-language pathologist present in the room during the play session took running notes of all utterances.

Transcripts from these 30-minute play sessions were prepared from the tapes, with every utterance and action of the children and mothers recorded using conventions established by the CHILDES consortium (MacWhinney, 1991). A variety of undergraduate and graduate student transcribers did the basic transcript preparation from audiotape, then checked all utterances on the videotape and added all actions, gestures, and relevant contextual information. The first author then checked the transcripts against the tapes and the running session notes that had been made by the speech-language pathologist. Every transcript was then checked against the tapes by the second author, a certified speech-language pathologist. This elaborate checking and rechecking procedure was necessary because of the poor intelligibility manifested by so many of the late talkers at follow-up.

When the transcripts had been fully checked and corrected, the CLAN procedures (MacWhinney, 1991) were used to identify a corpus of the first 100 complete utterances, after imitations, immediate self-repetitions, single-word "yes" or "no" responses to questions, memorized songs/rhymes, and unintelligible utterances had been eliminated from the corpus. The CLAN MLU program was then run on this 100-utterance corpus.

Finally, each 100-utterance corpus was coded using the IPSyn (Scarborough, 1990a). The IPSyn measures syntactic and morphological development in four areas: noun phrase, verb phrase, question/negation, and sentence structure. Points for each of the four domains are summed to yield a total IPSyn score. Fifty of the 55 transcripts used in this study were IPSyn-coded independently by the first and third authors. Mean interrater percent agreement across all IPSyn items was 96%, with agreement across the four IPSyn domains all above 95%.

In order to analyze MLU and IPSyn scores using a standardized metric, we converted all scores into z scores using the benchmark mean and *SD* values provided by Scarborough (1990a). For three of the late talkers, the 30-minute play session did not yield 100 scorable utterances. Therefore, MLU and IPSyn were calculated based on 75 utterances ($n = 1$) or 50 utterances ($n = 2$) for these children, and raw scores were prorated accordingly for calculation of IPSyn z score.

Group differences on the various outcome measures were analyzed using independent group t tests. Because of the number of t tests run, a p level of <.001 was used to determine statistical significance. To investigate the percentage of late talkers who were functioning in the average range on the various outcome measures, frequency distributions were computed using –1 *SD*, –1.5 *SD*s, and –2 *SD*s as cut-points. Finally, multiple regression analyses were run to examine predictors of age 3 outcome.

Results

Group Differences

Age 3 outcome data for the late talkers and comparison children are presented in Table 2. When followed up at age 3, children in the SLI-E group scored at the same level as comparison children in nonverbal ability on the Leiter International Performance Scale, consistent with their performance on the Bayley nonverbal items at intake. The late talkers were at age level on receptive language testing with the TACL–R, but the comparison children scored slightly higher, $t(37.85) = -2.61$, $p < .05$. This was consistent with receptive language performance at intake.

Table 2 also contains outcome scores for the two groups on the Expressive One-Word Picture Vocabulary Test (EOWPVT). As a group, the 28 late talkers tested

Table 2. Means and standard deviations of age 3 outcome measures by group.

Measure	Late talkers		Comparison children	
	M	*SD*	*M*	*SD*
Leiter Scale	131.67	15.29	134.29	15.29
TACL–R	98.63	7.88*	105.71	10.03
EOWPVT	99.50	19.92***	119.19	12.42
Reynell *z* score	–0.65	1.17***	1.48	00.86
MLU	2.46	0.94***	4.12	.64
MLU *z* score	–1.51	1.50***	1.13	1.02
IPSyn	48.26	18.48***	77.90	7.50
IPSyn *z* score	–2.21	1.56***	.40	.70
Noun phrase	–1.00	1.77***	.70	.54
Verb phrase	–1.84	1.59***	.64	.85
Question/Negs	–2.22	1.68***	.01	.88
Sent.Structure	–1.74	1.03***	–.03	.70

*$p < .05$, ***$p < .001$

on the EOWPVT scored in the average range, but significantly below the level of the comparison children with whom they had been matched at intake on SES and nonverbal cognitive ability, $t(45.68) = –4.24$, $p < .001$. On the Reynell Expressive Language Scale, although the late talkers scored well within 1 *SD* of age expectations at age 3, the comparison children scored more than 1 *SD* above age level, a significant difference, $t(50.73) = –7.68$, $p < .001$.

The most dramatic differences between SLI-E and comparison children at age 3 were in syntax. The late talkers at age 3 follow-up had an average MLU of less than 2.5, whereas the comparison children had a mean MLU of more than 4.0. As can be seen in Table 2, the late talkers had an MLU *z* score of –1.51, whereas the comparison children were more than 1 *SD* above age expectations, $t(52.04) = –7.91$, $p < .001$.

The IPSyn also showed dramatic differences between groups at age 3 follow-up (total score of 48.27 vs. 77.90). The groups also differed significantly in variance on the IPSyn, with the late talkers manifesting more than double the variance of the comparison children (18.48 vs. 7.50). As can be seen in Table 2, late talkers scored on average more than 2 *SD*s below age expectations on IPSyn total score, whereas the comparison children's *z* score was slightly above age level, $t(48.82) = –8.38$, $p < .001$. The data in Table 2 for the four IPSyn subscales make it evident that the late talkers were significantly different from comparison children in all four areas tapped by the scale, although they looked somewhat less delayed in noun phrase development and most delayed in questions/negations.

In sum, the general pattern of findings using group means at outcome is that the late talkers were significantly different from the comparison children on all expressive language measures. The late talkers as a group performed in the average range on the EOWPVT and the Reynell Expressive Language Scale, both measures of vocabulary, verbal concepts, and descriptive language use. In contrast, their mean scores were strikingly below age expectations on MLU and IPSyn, measures that tap various aspects of syntax, particularly inflectional morphology and grammatical structures.

Differences in Outcome Within the Late Talker Group

It is necessary to look beyond group means in order to address the issue of what percentage of former late talkers were performing in the average range on these four expressive language measures at age 3 outcome. For purposes of this analysis, we have defined five delay categories: none (performance within 1 *SD* of average), mild (performance ranging from –1 *SD* up to –1.5 *SD*s below average), moderate (delays of 1.5 to 2.00 *SD*s below age level), severe (delays of 2.00 *SD*s up to 3.0), and extreme (delays of 3 *SD*s or more). As shown in Table 3, the percentage of children performing in the average range varies considerably as a function of which expressive language measure is used.

As can be seen in Table 3, 79% the late talkers given the EOWPVT at 3 earned scores within 1 *SD* of average. Most of the other late talkers had "mild" delays, 7% had a "severe" delay, and none had an "extreme" delay of 3 *SD*s or more on the EOWPVT.

Using Reynell Expressive Language Scale *z* score as an outcome measure, 58% of the late talkers performed within 1 *SD* of age expectations. Most of the others had "mild" or "moderate" delays (1 to 2 *SD*s below age level), 9% scored 2 *SD*s or more below age norms, and none performed in the "extreme" delay range on the Reynell.

Table 3. Distribution of late talkers among expressive language delay outcome categories.

Measure	Delay categories[a]				
	None	Mild	Moderate	Severe	Extreme
EOWPVT	79%	11%	4%	7%	0%
Reynell	58%	21%	12%	9%	0%
MLU	35%	12%	9%	26%	18%
IPSyn	24%	18%	9%	18%	32%

[a]**Delay category criteria:**
None = within 1 *SD* of average performance at age 3.
Mild = between –1 *SD* and –1.49 *SD*s below age level.
Moderate = between –1.50 and 1.99 *SD*s below age level.
Severe = between –2.00 and 2.99 *SD*s below age level.
Extreme = –3.00 *SD*s or more below age level.

JSLHR, Volume 40, 556–566, June 1997

An even worse outcome is apparent when MLU z score is examined: 35% scored within 1 *SD* of age expectations, 21% scored between 1 and 2 *SD*s below average ("mild-to-moderate" delay), 26% were more than 2 *SD*s below age level, and 18% were performing more than 3 *SD*s below age expectations in MLU at age 3.

Finally, IPSyn total score showed the most impaired pattern of outcome, with only 24% of the late talkers scoring within 1 *SD* of age expectations. Of the rest, 27% scored 1 to 2 *SD*s below age level, 18% scored more than 2 *SD*s below average, and 32% scored 3 *SD*s or more below age expectations on the IPSyn ("extreme" delay).

In sum, depending on which of these four language measures is employed, the percentage of late talkers who performed in the average range (e.g., within 1 *SD* of age expectations) ranged from 79% on the EOWPVT (single word vocabulary) to 24% on the IPSyn (maturity of syntactic forms in noun phrase, verb phrase, questions/negations, and sentence structure). Correspondingly, the percentage of children delayed at least 2 *SD*s below age expectations ranged from 7% on the EOWPVT to 50% on the IPSyn.

This striking discrepancy across outcome measures seems mainly because the EOWPVT and the Reynell are primarily lexical measures, whereas MLU and the IPSyn tap syntax. It is also possible that the particularly favorable outcome picture obtained with the EOWPVT is due to its outdated and inflated norms; in fact, the EOWPVT was re-normed during the course of this follow-up study. Additionally, it should be noted that the six late talkers who did not receive the EOWPVT at age 3 had mean MLU and IPSyn z scores of –2.50 and –3.49 respectively. Thus, these six children, all early recruits into the SLI-E cohort, turned out to be among the most impaired youngsters in the group at the age 3 follow-up. Had they been omitted from the MLU and IPSyn outcome data, the percentage of children in the "severe-to-extreme" delay categories (>2 *SD*s below age expectations) would have dropped from 50% to 43% for IPSyn and from 44% to 36% for MLU. In short, this missing data factor contributed somewhat to the better outcome found on the EOWPVT versus the MLU and IPSyn, but the general pattern of findings is similar whether these six subjects are included or removed from the analysis.

Predictors of Age 3 Outcome

Despite the disparity in outcome status that emerged when lexical versus syntactic measures were used, these four measures of expressive language were quite highly and significantly intercorrelated within the late talker sample. MLU and IPSyn correlated at .86, consistent with other research (Scarborough, 1990a). The Reynell Expressive score correlated at .70 with both MLU and IPSyn, and the EOWPVT was significantly correlated with the other three outcomes measures, but at a somewhat lower level (correlations ranging from .44 to .63).

A central question motivating this study was whether or not age 3 expressive language outcome could be predicted from intake status at age 24 to 31 months. Our goal was to look for possible predictors of outcome in this sample of children with SLI-E, given that all the children in the sample had normal nonverbal ability and good receptive language. To this end, we carried out multiple regression analyses on our four major outcome variables (EOWPVT, Reynell, MLU, and IPSyn). The same four predictors were entered in the same order for each analysis. Intake expressive language delay was entered first (Reynell Expressive z score, based on age expectations), on the assumption that differences in intake expressive language delay severity would best predict age 3 outcome. Next, we entered intake receptive language delay (Reynell Receptive z score) and nonverbal cognitive ability (Bayley nonverbal item total score), to see if these intake variables would significantly improve prediction. Finally, we entered intake age, which had demonstrated a significant negative correlation with outcome in Rescorla and Schwartz (1990). Results of these regression analyses appear in Table 4. As can be seen in Table 4, the four regression analyses were quite consistent. The degree of delay in expressive language skill at intake using an age-normed score explained a significant percentage of the variance in age 3 expressive language outcome for all four outcome measures, with r-squares ranging from 21% for MLU to 34% for EOWPVT. Intake receptive language delay and intake nonverbal ability did not emerge as significant predictors of outcome for any of the four outcome measures. Finally, intake age was not a significant predictor of language outcome in any of the regressions, once the variance due to intake expressive language delay had been taken into account. It should be noted that intake age had negative correlations with all four outcome measures (ranging from –.50 with MLU to –.60 with IPSyn), but it was also highly correlated (r = –.76) with intake Reynell Expressive Language Scale z score, which represents each child's degree of lag below normative expectations using standard deviation units. Thus, Reynell Expressive z score and intake age are largely redundant measures in this study. Given the fact that most of the late talkers had expressive language skills in the same narrow range, the older the child was at time of diagnosis, the more delayed relative to age expectations he was and, therefore, the lower his z score would be.

Table 4. Regression analyses between intake and outcome measures.

Predictor	*R*-square	*R*-square increase	*F*	*p*
Hierarchical regression predicting age 3 EOWPVT				
Reynell exp *z*	.34	.34	13.38	.001
Reynell rec *z*	.38	.04	1.80	ns
Bayley nonverbal	.38	.00	.03	ns
Intake age	.42	.04	1.51	ns
Hierarchical regression predicting age 3 Reynell				
Reynell exp *z*	.25	.25	10.23	.01
Reynell rec *z*	.27	.03	1.11	ns
Bayley nonverbal	.30	.03	1.07	ns
Intake age	.33	.03	1.32	ns
Hierarchical regression predicting age 3 MLU				
Reynell exp *z*	.21	.21	8.67	.01
Reynell rec *z*	.21	.00	.00	ns
Bayley nonverbal	.23	.02	.63	ns
Intake age	.27	.04	1.62	ns
Hierarchical regression predicting age 3 IPSyn				
Reynell exp *z*	.33	.33	15.75	.001
Reynell rec *z*	.34	.01	.24	ns
Bayley nonverbal	.36	.03	1.41	ns
Intake age	.40	.04	1.89	ns

Discussion

The results of this study complement preliminary findings from this longitudinal project reported in Rescorla and Schwartz (1990). In the present study, which involved most of the late talkers in the project (34 out of the original 40) and which also employed a normally developing comparison group, children diagnosed with SLI-E between the ages of 24 and 31 months scored as a group in the average range on the EOWPVT and within 1 *SD* of age level on the Reynell Expressive Language Scale. However, on both these measures, the late talkers scored significantly below comparison children with whom they had been matched at intake on nonverbal ability and SES. At age 3, the late talkers were more than 1.5 *SD*s below age expectations in MLU and more than 2 *SD*s below expected levels on the IPSyn.

It would appear from these findings that these late talkers, who were identified primarily for delays in lexical development at 24 to 31 months, were most conspicuously delayed in syntactic complexity and morphological maturity by 36 months of age. This pattern of results is consistent with findings from the Portland sample (Paul, 1993), in which late talkers were much more likely to catch up in single-word vocabulary than in syntax. Our finding that the greatest delay at outcome was found on the IPSyn is also consistent with Paul and Alforde (1993), who reported that grammatical morphology was still delayed in late talkers who had caught up to comparison children in terms of MLU. However, unlike in the Portland sample, the late talkers in the present study were still significantly inferior to comparison peers in vocabulary skills, despite scoring in the normal range.

Examination of individual differences within the late talker group revealed that the proportion of children performing in the average range (within 1 *SD* of age level) at age 3 outcome varied widely as a function of expressive language measure used. If one takes the EOWPVT as the "gold standard" for outcome, 79% of the late talkers could be seen as performing at age level. This is consistent with results reported from the New York sample (Whitehurst et al., 1992), in which 88% of late talkers scored in the average range on the EOWPVT by age 3:6. A somewhat less positive outcome is indexed by the Reynell Expressive Language Scale, on which the children had to label pictures, define words, and describe pictures and were given a global rating of syntactic maturity. Using this general language measure, 58% of the late talkers could be said to have attained normal language status.

When measures of syntactic development (e.g., MLU and IPSyn) are taken as the "gold standard" of outcome, it becomes evident that most of our late talkers were still conspicuously delayed in their expressive language skills at age 3. Of our late talkers, only 35% were in the average range on MLU, and only 24% were at age level on the IPSyn when seen for follow-up at age 3.

Comparison of these syntax outcomes with those from previous studies suggests the importance of intake age in determining syntax outcome of young late talkers. For example, in the San Diego sample (Thal et al., 1991; Thal & Tobias, 1992), where the children ranged from 18 to 32 months at intake, 60% performed in the average range for MLU one year later (e.g., at age 30 to 42 months). In the Portland sample (Paul, 1993), with intake age ranging from 20 to 34 months, 40% of late talkers scored above the 10th percentile in DSS at age 3, and intake age was found to be a significant predictor of outcome. In the Wisconsin sample (Ellis Weismer et al., 1994), which had four children age 25 to 26 months at intake, only one child was within 1 *SD* of average in MLU at age 3 (25%). Thus, our figures of 35% within 1 *SD* for MLU and 24% within 1 *SD* for IPSyn would seem most consistent with results reported by Ellis Weismer, whose subjects, like ours, were at least 24 months old at intake. However, Ellis Weismer's oldest subject was 26 months old at intake, whereas the oldest children in our

JSLHR, Volume 40, 556–566, June 1997

sample of late talkers were 31 months of age. This may explain why only 25% (one child) of the Wisconsin sample was more than 2 *SD*s below age level in MLU, whereas 44% of the late talkers in our group performed at this level of severity.

Although the present report does not deal with outcome of these late talkers beyond age 3, longer-term outcome data for the Pennsylvania sample have been presented elsewhere (Rescorla, 1993; Rescorla & Dahlsgaard, 1995). These data indicate that although the late talkers in this study performed on average within age expectations on most language tests by age 5, they continued to have significantly weaker language skills through age 8 than the comparison children with whom they had been matched at intake.

Finally, the results of this study revealed that age 3 outcome could be significantly predicted by degree of expressive language delay at intake relative to age expectations (e.g., Reynell Expressive *z* score). As in Paul (1993), intake age was negatively correlated with outcome. Although our data indicate that intake age and intake expressive language delay were correlated with age 3 outcome at virtually the same level, the high degree of intercorrelation between these two predictors meant that the variance in outcome they accounted for was largely shared variance. Once intake expressive delay severity, an age-based variable, was taken into account, intake age became redundant as a predictor. Consistent with Ellis Weismer et al. (1994), Fischel et al. (1989), Paul (1993), and Paul, Spangle-Looney, and Dahm (1991), but inconsistent with Thal and Tobias (1992), intake receptive language skill was not predictive of outcome in this sample of late talkers, all of whom had language comprehension within the normal range for their age. Intake nonverbal ability also accounted for no significant variance in outcome, consistent with other studies of late-talking toddlers.

There are several important clinical implications from this study. All of our late talkers had more or less the same limited language skills (e.g., fewer than 50 words or no word combinations). However, this level of language is a more severe delay for a 30- month-old than it is for a 24-month-old (as indexed by an age-based score such as the Reynell *z*), because the period from 24 to 30 months is a time of very rapid progression in expressive language for normally developing children. In essence, the older a late talker is within this time period, the more he is falling behind on a steeply accelerating curve. Although children who are identified as SLI-E at 24 months may well be "late bloomers" who will look relatively normal by age 3, those diagnosed with SLI-E near 30 months of age have had a longer time to "bloom" but have not done so, suggesting that their underlying language deficiency may be more intractable and severe and, therefore, that their age 3 outcome will be worse.

Findings from the present study are of course limited in their generalizabilty. Children who also have receptive language delays and/or delays in nonverbal cognitive abilities are likely to have a worse outcome than the children in this sample, whose delay was in expressive language only. Furthermore, the children in this study generally came from privileged, middle to upper-middle class, intact families. Thus, these children with SLI-E were more likely to have had the kinds of input that are conducive to good language progress, including a variety of forms of early intervention services, than less-advantaged late talkers. Because we had no control over the provision of therapeutic services, we were not able to systematically examine their relationship to outcome as part of this research. However, it was our impression that the most impaired children were generally the most likely to be taken for speech-language treatment. Hence, without such intervention, outcomes might have been even worse.

Despite these caveats, our study suggests that parents and health practitioners need to be aware that a 24-month-old child with little or no expressive speech may be at significant risk for continuing language delay, even if he comes from a privileged family, has good nonverbal ability, manifests normal receptive language, and seems appropriately communicative. In fact, more than half such children are likely to manifest a significant delay in expressive syntax by age 3. Furthermore, such children perform generally below the level of comparison children with similar backgrounds and equal nonverbal skills in virtually every area of expressive language up through age 8, even though they are generally not still delayed in terms of normative expectations. This suggests that early expressive language delay may be a precursor of later subclinical individual differences in language skill, even when it does not presage long-term language disorder. Finally, the risk of continuing language delay appears to rise significantly within the period from 24 to 30 months of age. This suggests that it is appropriate to have mild concern for a 24-month-old who is not talking, but that this concern should become much more intense as the child passes the 30-month point and still has little or no expressive speech.

Acknowledgments

This research was supported by grants to the first author from the Bryn Mawr College Faculty Research Fund and from the National Institutes of Health (NICHD Area Grant 1-R15-HD22355-01; NIDCD R01-DC00807). The authors are indebted to Hollis Scarborough for her generous assistance in the area of IPSyn consultation and wish to thank the parents and children whose participation made this research possible.

References

Aram, D., Ekelman, B. L., & Nation, J. E. (1984). Preschoolers with language disorders: 10 years later. *Journal of Speech and Hearing Research, 27*, 232–244.

Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: The Psychological Corporation.

Bishop, D., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders, and reading retardation. *Journal of Child Psychology and Psychiatry, 30*, 1027–1050.

Bishop, D. V. M., & Edmundson, A. (1987). Language impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*, 156–173.

Carrow-Woolfolk, E. (1985). *Test of Auditory Comprehension of Language–Revised*. Allen, TX: DLM Teaching Resources.

Catts, H. W. (1991). Early identification of dyslexia: Evidence from a follow-up study of speech-language impaired children. *Annals of Dyslexia, 41*, 163–175.

Ellis Weismer, S., Murray-Branch, J., & Miller, J. F. (1994). A prospective longitudinal study of language development in late talkers. *Journal of Speech and Hearing Research, 37*, 852–867.

Fischel, J. E., Whitehurst, G. J., Caulfield, M. B., & Debaryshe, B. (1989). Language growth in children with expressive language delay. *Pediatrics, 82*, 218–27.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. S. (1993). *Guide and technical manual for the MacArthur Communicative Development Inventories*. San Diego, CA: Singular Press.

Fundudis, T., Kolvin, I., & Garside, R. F. (1979). A follow-up of speech retarded children. In C. A. Hersov & M. Berger (Eds.), Language and language disorder in childhood (book supplement to *Journal of Child Psychology and Psychiatry, 2*).

Gardner, M. F.(1981). *Expressive One-Word Picture Vocabulary Test*. Novato, CA: Academic Therapy Publications.

Klackenberg, G. (1980).What happens to children with delayed speech at 3? *Acta Paediatrica Scandinavia, 69*, 681–685.

Lee, L. (1974). *Developmental sentence analysis*. Evanston, IL.: Northwestern University Press.

MacWhinney, B. (1991). *The CHILDES language project: Tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum.

Miller, J. (1981). *Assessing language production in children: Experimental procedures*. Baltimore, MD: University Park Press.

Paul, R. (1991). Profiles of toddlers with slow expressive language development. *Topics in Language Disorders, 11*, 1–3.

Paul, R. (1993). Patterns of development in late talkers: Preschool years. *Journal of Childhood Communication Disorders, 15*, 7–14.

Paul, R., & Alforde, S. (1993). Grammatical morpheme acquisition in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech and Hearing Research, 36*, 1271–1275.

Paul, R., Laszio, C. M., McFarland, L. S., & Midford, N. (1993). Language outcomes in late-talkers: Kindergarten. *Society for Research in Child Development Abstracts, 9*, 534.

Paul, R., Spangle-Looney, S., & Dahm, P. S. (1991). Communication and socialization skills at ages 2 and 3 in "late talking" young children. *Journal of Speech and Hearing Research, 34*, 858–865.

Rescorla, L. (1989). The language development survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders, 54*, 587–599.

Rescorla, L. (1993, March). *Outcome of toddlers with specific expressive language delay (SELD) at ages 3, 4, 5, 6, 7, and 8*. Presented at Biennial Meeting of SRCD.

Rescorla, L., & Dahlsgaard, K. (1995, March). *Language and reading outcomes at ages 5-to-8 for toddlers with specific expressive language impairment*. Presented at Biennial Meeting of Society for Research in Child Development.

Rescorla, L., & Schwartz, E. (199O). Outcome of toddlers with expressive language delay. *Applied Psycholinguistics, 11*, 393–407.

Reynell, J. (1977). *Reynell Developmental Language Scales*. Windsor: NFER.

Richman, N., Stevenson, J., & Graham, P. (1982). *Preschool to school: A behavioral study*. New York: Academic Press.

Scarborough, H. S. (1990a). Index of Productive Syntax. *Applied Psycholinguistics, 11*, 1–12.

Scarborough, H. (1990b). Very early language deficits in dyslexic children. *Child Development, 61*, 1728–1743.

Scarborough, H. (1991). Early syntactic development of dyslexic children. *Annals of Dyslexia, 41*, 207–220.

Silva, P. A., Williams, S., & Mcgee, R. (1987). A longitudinal study of children with developmental language delay at age three: Later intelligence, reading, and behavior problems. *Developmental Medicine and Child Neurology, 29*, 630–640.

Sparrow, S., Balla, D., & Cicchetti, D. (1984). *Vineland Adaptive Behavior Scales*. Circle Pine, MN: American Guidance Service.

Tallal, P. (1988). Developmental language disorders. In J. F. Kavanaugh & T. J. Truss, Jr. (Eds.), *Learning disabilities: Proceedings of the national conference*. Parkton, MD: York Press.

Thal, D., & Bates, E. (1988). Language and gesture in late talkers. *Journal of Speech and Hearing Research, 31*, 115–123.

Thal, D., & Tobias, S. (1992). Communicative gestures in children with delayed onset of oral expressive vocabulary. *Journal of Speech and Hearing Research, 35*, 1281–1289.

Thal, D., Tobias, S., & Morrison, D. (1991). Language and gesture in late talkers: A 1-year follow-up. *Journal of Speech and Hearing Research, 34*, 604–612.

Tomblin, J. B., Freese, P. R., & Records, N. L. (1992). Diagnosing specific language impairment in adults for the purpose of pedigree analysis. *Journal of Speech and Hearing Research, 35*, 832–843.

Whitehurst, G. J., & Fischel, J. E. (1994). Early developmental language delay: What, if anything, should the clinician do about it? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 35*, 613–648.

Whitehurst, G. J., Fischel, J. E., Arnold, D. S., & Lonigan, C. J. (1992). Evaluating outcomes with children with expressive language delay. In S. F. Warren & J. Reichle (Eds.), *Causes and effects in communication and language intervention, Vol. 1* (pp. 277–313). Baltimore, MD: Brookes Publishing.

Whitehurst, G., Smith, M., Fischel. J., Arnold, D., & Lonigan, C. (1991). The continuity of babble and speech in children with specific expressive language delay. *Journal of Speech and Hearing Research, 34*, 1121–1129.

Received March 2, 1994

Accepted December 13, 1996

Contact author: Leslie Rescorla, PhD, Department of Psychology, 101 N. Merion Avenue, Bryn Mawr, PA 19010. Email: lrescorl@brynmawr.edu