INT. J. LANG. COMM. DIS., JANUARY–MARCH 2005,
VOL. 40, NO. 1, 1–27

Speech, language and literacy skills 3 years later: a follow-up study of early phonological and metaphonological intervention

Barbara Bernhardt and Eva Major

School of Audiology and Speech Sciences, Vancouver, Canada

(*Received 3 May 2002; accepted 20 February 2004*)

Abstract

Background: Three years before the present study, 19 preschool children participated in a phonological and metaphonological intervention programme. The phonological intervention programme was based on non-linear phonological analyses. The metaphonological intervention programme included both rhyming and alliteration tasks and was directly targeted during the last section of the programme. All children made significant gains in phonology, and many in metaphonology. However, the literature notes a tendency for children with history of early speech or language impairments to have difficulty acquiring literacy skills or to show residual speech impairment. The participants of the 1998 study were therefore considered at risk for continuing speech impairment and/or academic performance.
Aims: The primary objectives were to document the children's later speech, language and literacy skills, and to determine potential relationships between previous and concurrent child factors.
Methods & Procedures: Twelve children from the original cohort (aged 6;1–8;5) received a comprehensive battery of speech, language, cognitive and academic tasks 3 years after the original study. Data were evaluated both concurrently and in terms of the children's preschool performance.
Outcomes & Results: Five of the children had residual phonological impairment. Only two children showed below average reading (decoding and comprehension), although five also showed below average spelling performance. Children with limitations in verbal memory, language production and metaphonology at the follow-up point were more likely to show delays in literacy. Language production skills at the follow-up point were most strongly correlated with literacy development. In terms of the early study, the strongest predictor for literacy development was performance on metaphonology tasks at the end of the early intervention study. The strongest predictor for ongoing speech impairment was phonological skill at the end of the early study.

Address correspondence to: Barbara Bernhardt, School of Audiology and Speech Sciences, 5804 Fairview Avenue, Vancouver BC, Canada V6T 1Z3; e-mail: bernharb@interchange.ubc.ca

International Journal of Language & Communication Disorders
ISSN 1368-2822 print/ISSN 1460-6984 online © 2005 Royal College of Speech & Language Therapists
http://www.tandf.co.uk/journals
DOI: 10.1080/13682820410001686004

Conclusions: Results suggest that early phonological and metaphonological intervention can promote normalization of speech development and normal acquisition of literacy skills for at least some children with a history of severe phonological impairment. In the earlier study, the component structure of words (onsets, rhymes, codas) was emphasized through both metaphonological and non-linear phonological intervention. Risk for literacy and ongoing speech impairment can be reduced through early intervention that draws attention to the structure of words.

Keywords: phonological impairment, metaphonological intervention, follow-up, literacy.

Introduction

The immediate goal of phonological intervention is to facilitate acceleration of speech development. The expectation is that children's phonological skills will be like those of peers with similar cognitive abilities after intervention. A broader goal of phonological intervention is the enhancement of language and literacy skills. Although the first goal is often met (Gierut 1998), studies have suggested that children with early history of phonological impairment may be at risk for continuing speech impairment (Shriberg *et al.* 1994) and/or delays in acquisition of literacy skills (e.g. Magnusson and Naucler 1989, 1993, 1998, Bishop and Adams 1990, Catts 1993, Clarke-Klein 1994, Gillon and Dodd 1994, 1995, Bird *et al.* 1995, Webster *et al.* 1997, Gillon 2002). Major and Bernhardt (1998) reported immediate and positive outcomes of a phonological and metaphonological intervention study with 19 preschool children. Because the literature suggested that some of these children might show continuing speech impairment and/or delays in later literacy skill acquisition, a follow-up study was conducted 3 years after the end of the earlier study.

Early phonological impairment and later development

Given the current level of knowledge about phonological impairment in children (defined broadly as speech sound impairment of unknown origin), it is not possible to predict which children with early phonological impairments will 'normalize' quickly and which will have ongoing speech and/or language/literacy impairments. Shriberg and colleagues differentiate between children whose speech 'normalizes' by age 6 (their term, short-term 'normalization'), and those whose speech normalizes after that point. Children in their studies showed short-term normalization typically about 1–2 years after identification and treatment, and long-term normalization, about 5 years after identification and treatment (Shriberg *et al.* 1994a, b). Not all children in their studies achieved normal speech within 5 years; some showed plateaus in speech development after age 8 (Shriberg *et al.* 1994b). In Webster *et al.*'s (1997) prospective study, about two-thirds of their 29 participants showed short-term normalization. The average pretreatment per cent consonants correct score for children in their sample was 50.6% (SD 39.2%), or a moderate severity level, according to Shriberg and Kwiatkowski (1982). This suggests that children with more moderate impairments may show short-term normalization. Shriberg *et al.* (1994b) and Hesketh *et al.* (2000) also found severity to be significantly correlated with rate of change in intervention, but noted variability across children. In trying to determine why some children show faster normalization rates, Shriberg *et al.* (1994a, b) examined many factors (language performance, cognition, hearing, oral motor factors, treatment time, parent involvement, motivation and focus). Only two variables distinguished between rates of phonological normalization in their samples. The first was phonological: children showing fewer deletions (reductions of word structure) tended to normalize more quickly. The second was psychological: children who were more motivated and focused tended to show faster normalization. These studies suggest four main factors to consider when examining patterns of normalization: age of the child, time in therapy, type and severity of early impairment (deletions versus substitutions), and degree of motivation and focus. Another factor related to rate of normalization might be metaphonological awareness skills pretreatment. In terms of metaphonological awareness, Hesketh *et al.* (2000) noted that pretreatment metaphonology was somewhat predictive of outcomes in phonological intervention; children with better-developed phonological awareness skills pretreatment showed a faster rate of change in their study but this was not a significant trend. Overall, much more research is needed in the area of prediction of speech normalization rate.

Comparatively speaking, more outcome studies have been conducted in recent years on the literacy skill development of children with a history of early phonological impairment. Reading and writing are complex tasks, and related to many child-internal and external variables, e.g. age, severity of impairment, general language ability, speech perception, phonological processing skills, intelligence, family and environmental factors. First, in terms of age, Naucler and Magnusson (1999) noted that children under age 8 tended to show difficulty with a variety of decoding, reading comprehension, and spelling tasks, but by the end of the primary grades, only with reading comprehension. Then, by the end of Grade 12, adolescents in their sample who continued to show a learning difficulty again showed depressed scores on all tasks: spelling, word and text comprehension (Naucler and Magnusson 1999). The age of speech normalization may also be relevant. Children who have adequate phonological skills at the time they start to read and write may be less likely to demonstrate difficulty acquiring literacy skills than those who continue to have phonological deficits (Bishop and Adams 1990, Webster *et al.* 1997, Gillon 2002). In a similar vein, severity of phonological impairment (sometimes indicative of long-term normalization) has been positively correlated with reading disability in some studies (Magnusson and Naucler 1989, Webster and Plante 1992, Bird *et al.* 1995), although Magnusson and Naucler (1989) did observe individual difference in this regard. Children with what they called 'syntagmatic disorders' (involving assimilations and metatheses) were more likely to show later difficulties with literacy tasks than children with more 'paradigmatic disorders' (involving deletions and substitutions of individual speech sounds).

General language ability has been implicated in several studies (Shriberg and Kwiatkowski 1988, Bishop and Adams 1990, Lewis and Freebairn 1992, Catts 1993, Johnson *et al.* 1999). Children with language production and comprehension impairments in addition to their phonological impairments were more likely to have reading disabilities than children with phonological impairments only. Shriberg *et al.* (1994a) suggest that 50–75% of children with speech delay have a delay in language production, and 10–40% an additional delay in language comprehension. Thus, at least half of the children with early phonological impairment might be expected to have later difficulty with literacy skill acquisition. This estimate is consistent with Catts *et al.*'s (1999) observation that over 50% of children in their samples with history of early impairments had later reading disabilities.

Other child-internal variables that have shown positive correlations with literacy skills include speech perception (Brady *et al.* 1983, McBride-Chang 1995), intelligence (Shriberg and Kwiatkowski 1988, Bishop and Adams 1990, Bird *et al.* 1995) and various phonological processing skills. A variety of phonological processing tasks have shown positive correlations with literacy skills, most commonly, phonological working memory and metaphonological awareness tasks, but also rapid naming (Snowling *et al.* 1988, Catts *et al.* 1999). Phonological working memory has been tested in a variety of ways, using sentence imitation (Webster and Plante 1992, Webster *et al.* 1997, Catts *et al.* 1999), non-word imitation (Gathercole and Baddeley 1993, Gillon and Dodd 1995, Catts *et al.* 1999) or other imitative recall procedure (Fazio 1986, Gillam and Van Kleeck 1996). Metaphonological awareness or 'phonological sensitivity' has been tested by a variety of tasks such as rhyming, alliteration, or segmentation of various units, with or without picture support, and through detection and/or production tasks (Magnusson and Naucler 1989, 1993, Catts 1993, Bird *et al.* 1995, Webster *et al.* 1997, Gillon 2002).

In addition to these child-internal variables, family, educational, and environmental factors have also been shown to correlate strongly with literacy outcomes for children (Snow *et al.* 1998).

The observed relationships between literacy skills and other factors have varied considerably across studies because of considerable methodological differences in definitions, tasks, analyses and sampling. In terms of definitions, phonological impairment (or disorder) has several connotations. Some researchers differentiate between phonological delay, disorder, and various motor speech disorders (e.g. Dodd 1995). Others define phonological impairment more broadly as a developmental speech impairment of unknown origin (Stoel-Gammon and Dunn 1985, Shriberg and Kwiatkowski 1994, Bernthal and Bankson 1998, Bernhardt and Major, present study). Severity of phonological impairment is also defined differently across studies, in terms of degree of deviation from the norm on a standardized test, percentage of consonants correct (PCC, as in Shriberg and Kwiatkowski 1982), number and type of phonological processes (Khan and Lewis 1986, as in Webster *et al.* 1997, Magnusson and Naucler 1989), or proportion of word structure reductions (Bernhardt and Stemberger 2001). There have been ongoing debates in the literature about metaphonological awareness, as to whether it reflects one underlying ability (Anthony *et al.* 2002), or whether the various skills reflect independent abilities (Muter and Snowling 1998).

In terms of measurement, there are many standard and non-standard ways to evaluate each of the variables listed above, and no consensus as to the optimal measure for any of them. One of the most highly studied variables in the past two decades is phonological sensitivity. Researchers have debated about the relative merit of rhyming tasks (e.g. Bryant 2002, Goswami 2002) versus phoneme awareness tasks (Hulme *et al.* 2002) as predictors of literacy skill development. There are also many ways to evaluate language, speech perception, phonological working memory, and intelligence. Depending on the method used, the researcher may or may not find significant positive correlations with literacy skill acquisition.

Sampling size and type has also affected results in this area. Most studies investigating the relationship between phonological impairments and other skills have been with small numbers of participants. In such studies, individual difference can skew results, giving weight to one variable or another. In a larger *n* study, Catts *et al.* (1999) found *both* oral language skills and phonological processing skills to be correlated with literacy skills across participants. Whitehurst and Lonigan (1998) also drew that conclusion in a general review of the literature regarding literacy development. Results in the larger *n* studies confirm the findings of smaller *n* studies, that the relationship between early phonological impairment and later literacy skill acquisition is complex and variable across individuals. Overall, studies indicate that general language performance, severity of phonological impairment, and phonological processing skills are highly correlated with literacy development. The particular significance of each of these variables for a given individual remains difficult to determine, however, given the conflicting methodologies and results across studies, and the influence of motivational factors and environmental supports for learning. At this point in time, there are many studies showing the *benefits* of phonological awareness (or metaphonological) training in the late preschool–early primary grades for literacy development (for a review, see Troia 1999) although, as yet, there are few such intervention studies concerning children with spoken language impairment (exceptions being Hesketh *et al.* 2000, Gillon 2002).

Overview of the earlier intervention study

Major and Bernhardt (1998) reported immediate phonological and metaphonological intervention outcomes for 19 preschool children with moderate to severe phonological impairments. These children participated individually in a 16-week intervention study based on non-linear phonological analyses and treatment approaches that emphasized the componential structure of speech sounds, syllables and words (see appendix). For the study, the data were collected by community speech–language pathologists who conducted the phonological intervention. Profiles are reported below for the 12 children who participated in the current follow-up study.

Participants' pre-intervention profiles (table 1)

At the beginning of the preschool intervention study, the 12 children participating in the follow-up study were between 3;3 and 4;11 years of age (mean 3;11). Socio-economic status of the participating families was in the low to middle range.

The children had normal hearing and oral-motor function at the time of testing. Language comprehension scores were also within normal limits, as measured by the *Peabody Picture Vocabulary Test—Revised* (PPVT-R; Dunn and Dunn 1981), and either the *Test of Auditory Comprehension of Language—Revised* (TACL-R; Carrow-Woolfolk 1985), the *Preschool Language Scale—3* (PLS-3; Zimmerman *et al.* 1992), or the *Reynell Developmental Language Scales* (RDLS; Reynell 1985).

Delays were noted on a number of tasks. A standard word list (Bernhardt 1990) was used to elicit a minimum of 160 spontaneous single words from each participant; this was supplemented with words from the children's connected speech samples. Eleven of the participants had severe phonological impairments pretreatment. PCCs for single words ranged from 10.9% (Colin) to 53.6% (Miles), with an average of 25.4%. (Pseudonyms are used throughout the present paper that match

Table 1. Pre- and post-intervention results in the preschool intervention study (Major and Bernhardt 1998)

Child pseudonym	Age (years;months) Pre-	Age (years;months) Post-	PCC Pre-	PCC Post-	Word shape match Pre-	Word shape match Post-	Metaphonology (/700) Pre-	Metaphonology (/700) Post-	IPSyn (1990) Pre-	IPSyn (1990) Post-
Jeanie	3;6	3;9	23.8	32.18	24.6	30.28	120	40	35	52
Marcy	3;6	3;9	16.99	25.47	20.79	29.22	0	20	21	36
Ben	4;0	4;7	26.32	50.00	17.09	47.49	20	140	35	55
Faith	3;6	4;0	33.41	42.62	44.44	64.43	153	284	77+	77+
Dan	3;8	4;3	22.96	44.91	26.32	57.69	103	165	19	49
Lloyd	3;6	3;11	21.54	25.83	26.67	39.66	20	0	45	58
Gary	4;4	5;0	22.68	38.66	25.10	52.94	87	396	66	66
Brad	3;6	4;1	22.66	40.97	20.90	47.54	0	40	15	56
Miles	4;0	4;6	53.60	60.17	48.08	69.75	133	50	55	79+
Serena	4;5	4;10	28.01	42.14	25.74	59.89	110	495	62	70+
Terry	4;7	5;0	21.75	25.65	27.85	38.69	273	510	43	66
Colin	4;11	5;4	10.90	34.96	19.11	48.26	40	313	15	55
Mean (SD)	3;7 (6.3 months)	4.5 (6.5 months)	25.4 (10.4)	38.6 (10.6)	27.2 (9.5)	48.8 (12.9)	88.2 (79.3)	204.4 (188.8)	40.7 (20.1)	54.8 (9.1)

PCC, per cent consonants correct for single words (including clusters) using the Bernhardt (1990) single word list.
Word shape match is the percentage of child CV sequences in words matching those in adult words. All children were delayed with respect to phonology, and most with respect to the IPSyn (Scarborough 1990). Metaphonology test tasks included nursery rhymes, rhyme and alliteration production, sentence segmentation, disyllabic and monosyllabic word segmentation, and word reconstruction (changing names of objects).
+Score within or above normal limits.

the pseudonyms of the earlier study.) According to the *Index of Productive Syntax* (IPSyn; Scarborough 1990), language production skills were also delayed for all but one participant (Faith). Among the twelve children, four children had no or few responses on the initial assessment of metaphonology. Metaphonology tasks included nursery rhymes, rhyme and alliteration production, sentence segmentation, disyllabic and monosyllabic word segmentation, and word restructuring (changing names, e.g. *me* into *bee*). The metaphonology tasks for the preschool study were based on the ones in Green (1991) that showed significant differences between preschool children with typical versus delayed phonological development. Although these tasks were not norm-referenced, children in the current cohort performed significantly poorer than typically developing 3-year-olds on the tasks (Bernhardt *et al.* 1995).

Participants' immediate post-intervention profiles

Testing at the end of the 16-week preschool study showed significant gains on all measures across participants. In terms of phonological development, there were significant gains in consonant and vowel production and in word shape matches (CV sequences) with adult targets. The average gain in PCC for single words was 13.3% across the study. The least improvement in PCC was noted for one of the youngest participants (Lloyd) and one of the oldest participants (Terry), each showing less than 5% gain. The average gain in word shape match across the study was 21.6%. The least improvement in word shape match was noted for the youngest participants, Jeanie and Marcy (less than 10%). Language production scores increased significantly over the study (Bopp 1995), although only three children had average range IPSyn (Scarborough 1990) scores (Faith, Miles, Serena). All but one young participant (Lloyd) showed improvement in metaphonological awareness across the study.

Purposes and predictions of the present study

The literature suggested that the participants of the 1998 study were at risk for continuing speech impairment and/or delays in later literacy skill acquisition. Thus, a follow-up study was conducted 3 years later, when all children were at least six years of age. At this point, all had reached the cut-off point for short-term normalization according to Shriberg *et al.* (1994a, b). The primary objectives of the follow-up study were (1) to document the children's later speech, language and literacy skills and (2) to determine potential relationships between previous and concurrent child factors. The longitudinal data also provide an indirect way to evaluate the early intervention, although that was not the primary purpose of this study. The non-linear phonological intervention emphasized the structure of words during intervention, as did the metaphonological training. A focus on word structure may have enhanced possible longer-term outcomes, as the following comment from Gillon (2002: 397) implies:

> Teaching children to segment words into phonemes, to focus on consciously perceiving individual phonemes in words and increasing the child's ability to hold information about a word's phonological structure in memory (such as required in a phoneme manipulation task) may help children to form more complete phonological representations of words. This may lead to improvements in speech production, and in decoding or encoding.

The children's pre- and post-intervention profiles led to a number of predictions for their speech and literacy skills at the follow-up point, with respect to the strong predictors from the literature: severity of phonological impairment, general language performance, and metaphonological abilities. All children were considered at risk for delayed literacy development because of at least one of those predictors.

- Severity of phonological impairment: 11 children were predicted to be at risk for ongoing speech impairment and delayed literacy development because of their severe phonological impairments pre-intervention (all but Miles). Colin, who was 5;0 pretreatment and had a PCC of 10.9%, was considered most at risk. The average pre-intervention PCC of the children in the Major and Bernhardt (1998) study was much lower than that of a similar longitudinal study (Webster *et al.* 1997): 25.4 compared with 50.9%. The data from the present study thus provided an opportunity to compare outcomes across studies with respect to severity of impairment pre-intervention. Those children who continued to show PCC scores of less than 40% after intervention were considered to be at greater risk for continuing speech impairment (Colin, Terry, Gary, Lloyd, Marcy, Jeanie). In that group, Colin and Terry were considered most at risk, because they were 5 years old at the end of the study.

- Language production delays: pre-intervention IpSyn (Scarborough 1990) scores suggested that all but Faith were at risk for delayed literacy development. After intervention, all but Faith, Miles, and Serena continued to be at risk based on IpSyn scores.
- Metaphonology: in terms of pre-intervention metaphonology scores, the four children at greatest risk for delayed literacy development were Marcy, Brad, Ben, Colin. Post-intervention, the five children who showed slow progress in metaphonology (Jeanie, Marcy, Lloyd, Miles, and Brad) were considered most at risk, although the two 4-year-olds (Miles and Brad) more so than the 3-year-olds.

In terms of the follow-up testing, it was predicted that children who continued to show low scores on language production, metaphonology, or phonology tasks would be more likely to show low scores on academic tasks. In addition, it was predicted that those who had low scores on verbal memory and/or non-verbal tasks would show delayed literacy development.

Methods

General study design

In terms of research methodology, this study falls into the descriptive, case study category, similar to many other studies in speech–language pathology (Olswang 1998). The follow-up investigation took place about 3 years after the preschool intervention study. Children were tested in their homes by one of the experimenters (the second author) over two to three 1.5-hour sessions (table 2) (see the following section for a description of the tasks).

Most tasks were scored during the assessment. For those tasks requiring transcription, audiotaped speech samples were collected using a Marantz PMD430 tape recorder and PMZ table-top microphone. The examiner transcribed the tapes first, and then the first author transcribed all multisyllabic words and 10% of the shorter words. Because the children had very intelligible speech, transcription agreement was over 95% for phones and diacritics. Where disagreement arose, a consensus transcription was created, with the authors listening to the tapes together and agreeing on segments or diacritics through discussion.

Data from the various tasks were statistically analysed using SPSS 11.0 (2001). Because previous research suggests that concurrent and post-intervention performance may be more highly correlated with literacy than pre-intervention performance (Bird *et al.* 1995, Webster *et al.* 1997), data from three testing points (pre-intervention, post-intervention, and concurrent) were compared with the follow-up literacy results.

General participant characteristics at the follow-up point

Four girls and eight boys participated in the follow-up study. (The other seven participants from the early study were not able to participate. Mann–Whitney *U*-tests revealed no significant differences between the 12 participants and the seven

Table 2. Assessment tools for the follow-up study

Domain	Tool
Phonology	*Goldman–Fristoe Test of Articulation—Revised* (GFTA-R, Goldman and Fristoe 1986) *Assessment of Phonological Processes—Revised* (APP-R, Hodson 1986) multisyllabic words
Word discrimination	*Test of Language Development—2 Primary* (TOLD-P 2, Newcomer and Hammill 1988): Word Discrimination subtest
Metaphonology	Non-standard tasks: rhyming and alliteration production, initial and final consonant deletion
Language comprehension	*Peabody Picture Vocabulary Test—Revised* (PPVT-R, Dunn and Dunn 1981) *Clinical Evaluation of Language Fundamentals—3* (CELF-3, Semel, Wiig and Secord 1995): Sentence Structure subtest
Language production	CELF-3 (1995): Word Structure and Word Association subtests Non-standard narrative
Verbal memory	CELF-3 (1995): Recalling Sentences subtest *Wechsler Intelligence Scales for Children—III* (WISC-III, Wechsler 1989) digit span subtests
Non-verbal skills	*Test of Nonverbal Intelligence—2* (TONI-2, Brown *et al.* 1990): Geometric Categories, Sequences, and Analogies subtests
Reading	*Peabody Individual Achievement Test—Revised* (PIAT, Dunn and Markwardt 1989): Reading Recognition and Reading Comprehension subtests
Spelling	PIAT-R (1989): Spelling subtest
Arithmetic	PIAT-R (1989): Arithmetic subtest

Table 3. Phonology and phoneme discrimination scores at the follow-up point

Child	Age (years;months)	Percentile ranks (and PCC) of GFTA-R (1986)[a]	PCC of Hodson (1986) multisyllabic words	TOLD-2 P (1988) percentiles, Word Discrimination
Jeanie	6;1	21 (88.7)*	79.7	25*
Marcy	6;2	19 (82.8)*	66.7	16*
Ben	6;10	27 (82.0)	55.6	63
Faith	6;11	60 (96.9)	94.6	63
Dan	7;1	2 (63.6)[b]**	64.2	16*
Lloyd	7;1	100 (99.2)$^+$	92.1	63
Gary	7;2	56 (96.5)	85.7	37
Brad	7;4	84 (96.2)	85.7	16*
Miles	7;5	8 (79.3)**	66.7	91$^+$
Serena	7;7	4 (81.9)**	89.7	63
Terry	8;3	2 (65.2)**	64.4	25*
Colin	8;5	>1 (63)**	74.5	16*
Mean (SD)	86.3 (8.4)	82.9 (13.3)	76.6 (12.99)	41.2 (26.04)

[a]Percentile ranks from the *Goldman–Fristoe Test of Articulation—Revised* (Goldman and Fristoe 1986). PCC, per cent consonants correct for single words.
[b]Score designation: *16–25th percentile; **greater than −1 SD; $^+$greater than +1 SD.

non-participants children in terms of pre- and post-intervention phonology and metaphonology scores.) Ages at the follow-up point ranged from 6;1 (Jeanie) to 8;5 (Colin), with a mean age of 7;2 (SD 8 months) (table 3).

Grades completed at the follow-up point ranged from kindergarten to Grade 2 (table 3: kindergarten is the first year of schooling and focuses on pre-literacy tasks and play; Grade 2 is the third year of schooling and the second academic year.). In the years following the early intervention study, all but three participants (Jeanie, Miles and Serena) had received some intervening speech–language services. (Details on the amount and type of intervention are unavailable.)

Follow-up study tasks (table 2)

Tasks used to evaluate the children's language and academic performance at the follow-up point were based on findings of other investigators, the availability of materials, and the time available for testing. Areas assessed were phonology, phonemic discrimination, language production and comprehension, metaphonology, verbal memory, non-verbal intelligence, and academic skills of reading (decoding and comprehension), spelling and arithmetic. Because no control group was used, standardized tests served as norm references where available.

Phonology

One of the objectives of the study was to determine whether the children's phonological skills were within normal limits. The *Goldman–Fristoe Test of Articulation—Revised* (GFTA-R; Goldman and Fristoe 1986) was used to compare the children's speech skills to developmental norms. In addition, the multisyllabic words from the *Assessment of Phonological Processes—Revised* (APP-R; Hodson 1986) were elicited. Those words were probably above the age level of many children at the follow-up point but provided a way to evaluate a child's ability to imitate complex word shapes and consonantal sequences. In this way, the multisyllabic word task was similar to non-word repetition tasks frequently used in other investigations. Non-word repetition tasks were not included, because of the inherent difficulty in creating non-words that are balanced in terms of phonemic and phonotactic frequency and familiarity.

Phonemic discrimination

Because speech perception ability has been correlated with literacy skills in some studies (Brady *et al.* 1983, McBride-Chang 1995), a norm-referenced task involving phonemic discrimination was included—the word discrimination subtest of the *Test of Language Development—2 Primary* (Newcomer and Hammill 1988). This is an orally presented task that requires same-different discrimination of words differing in one phoneme (e.g. *weak-weep*, same or different?).

Language production and comprehension

Studies have suggested that children who have delays in language development in addition to phonological delay may be at greater risk for literacy development (e.g. Bishop and Adams 1990, Lewis and Freebairn 1992, Catts 1993). Most children in the earlier study had shown delays in language production (mild to severe). In addition, Miles had shown borderline average language comprehension at the outset. Thus, it was considered important to assess general language performance.

The *Peabody Picture Vocabulary Test—Revised* (PPVT-R; Dunn and Dunn 1981) is used in many research studies, and had been administered to the children in this study when they were preschoolers; thus, it was one measure of choice. The *Clinical Evaluation of Language Fundamentals—3* (CELF-3; Semel *et al.* 1995) is a norm-referenced test of language function suitable for school-aged children. Selected subtests of language comprehension (Sentence Structure) and production (Word Structure, a morphological analogy test, and Word Association, within-category naming) were chosen. In addition, a narrative sample was elicited (although results are not reported here because of sampling inequities across children).

Metaphonology

Metaphonology tasks often show positive correlations with literacy tasks (e.g. Magnusson and Naucler 1989, 1993, Catts 1993, Bird *et al.* 1995, Webster *et al.* 1997, Gottardo *et al.* 1996, Gillon 2002). Participants in the follow-up study had shown a range of performance on metaphonology tasks in the earlier study, both before and after intervention. Thus, metaphonology was considered a critical variable to assess in this study. Because no commercially available test sufficiently covered the desired range of tasks across the age range of the participants, the following non-standardized metaphonological tasks were administered:

- Rhyme production: the child was presented orally with five different CVC words (one at a time), and was asked to produce three rhyming words per stimulus item (one point per item). No context was provided, although one example was given. Non-words were acceptable responses. Focus was on the vowel and final consonant. Stimuli were *bat*, *key*, *ball*, *dad*, and *toe*.
- Alliteration production: the child was asked to give three words beginning with the same sound as each of five orally presented monosyllabic words starting with single consonants (one point per item). One example was given. Non-words were acceptable responses. Focus was on the initial consonant only. Stimuli were *bee*, *fight*, *sock*, *win*, and *hat*.
- Initial consonant deletion: the child was presented with five CCVC words orally (one at a time) and asked to say each word again without the first sound. Three points were given for a phonemic deletion (*nail* for *snail*), two points for an onset–rime division (*ail* for *snail* and the /s/ named as the first sound) and one point for either an onset–rime division or naming of the first sound. Stimuli were *stop*, *snail*, *trip*, *broom*, and *flip*.
- Final consonant deletion: the child was presented with five CVC or CVCC words orally (one at a time) and asked to say the word again without the last sound. Three points were given for a phonemic deletion (*plan* for *plant*), two for an onset–rime division with the last consonant named (*une* for *tune*, and /n/ named), and one for either naming the last sound or producing an onset–rime division). Stimuli were *beak*, *soap*, *tune*, *plant*, and *tile*.

The alliteration and rhyming tasks were identical to those used in the earlier study. Many participants had found those tasks challenging at earlier ages, and thus it was of interest to determine whether they had mastered them. Because most children had developed some segmentation skills by the end of the first study, more difficult segmentation tasks were used in the follow-up study than in the early study in order to avoid ceiling responses across participants.

Verbal memory

Verbal memory skills have often been correlated with literacy skills (Fazio 1986, Gathercole and Baddeley 1993, Gillam and Van Kleeck 1996, Webster *et al.* 1997, Catts *et al.* 1999, Chiappe and Siegel 2000). The literature shows divergent results, however, depending on the tasks used. Thus, two types of recall tasks were used: the digit span subtest of the *Wechsler Preschool and Primary Intelligence Scale—III, Revised* (WISC-III; Wechsler 1989), and the Recalling Sentences subtest of the CELF-3 (Semel *et al.* 1995).

Non-verbal skills

In at least one previous study, non-verbal skills were significantly correlated with reading disability in children with phonological impairment (Bird *et al.* 1995). Three subtests from the *Test of Nonverbal Intelligence—2* (TONI-2; Brown *et al.* 1990) were administered: Geometric Categories, Geometric Sequences and Geometric Analogies. These tasks were considered to be least likely to engage verbal mediation.

Academic skills

The *Peabody Individual Achievement Test—Revised* (PIAT-R; Dunn and Marquardt 1989) was used to assess reading recognition and reading comprehension, spelling, and arithmetic. PIAT-R could be used across the age range of participants in this study and is a commonly used test of school achievement. The grade level of the younger participants (end of kindergarten) precluded any test that involved writing, which would have been required for many of the other commonly used academic tests. PIAT-R covered all domains of interest (reading recognition, reading comprehension, spelling, and arithmetic).

Results

Results are presented descriptively for each of the task sets, and then in terms of statistical analyses between early and follow-up variables. In the descriptive results, performance is compared with reference norms where such were available. Scores more than one standard deviation below the mean are designated as below average; scores between the 16th and 25th percentiles are designated as low average. Individual names are listed in descending order of scores.

Phonology and phonemic discrimination (table 3)

All children were intelligible in conversation. According to the GFTA-R (Goldman and Fristoe 1986) percentile ranking, 7/12 had average or above average scores. PCCs for the words elicited with the GFTA-R (1986) ranged from 63 to 99.2%

(mean 82.9%, SD 13.3%), i.e. a mild-to-moderate range of severity (Shriberg and Kwiatkowski 1982). Most non-matches with adult targets involved later-acquired phonemes: one or more of the liquids /r/ or /l/, coronal fricatives or affricates, or /v/. PCCs were also calculated for the Hodson (1986) multisyllabic words. These ranged from 55.6 to 94.6% (mean 76.6%, SD 13.99%).

All participants scored within normal limits on the Word Discrimination subtest of the TOLD-2P (Newcomer and Hammill 1988), although 6/12 children scored in the low average range.

The children whose scores were below the 25th percentile for phonology were Jeanie, Marcy (both low average), Miles, Serena, Dan, Terry, and Colin (all below average). Children showing low average performance in word discrimination were Terry, Jeanie, Brad, Marcy, Dan, and Colin.

Metaphonology, language production and comprehension (table 4)

Total metaphonology scores ranged from 18 to a ceiling score of 60. Children scoring below one standard deviation below the mean for the group were Ben, Marcy, and Miles. All but two participants gained points on each of the tasks. Marcy did not produce any rhymes, and Miles scored no points on the segmentation tasks.

In terms of language comprehension, children scored within normal limits or above on the Sentence Structure subtest of the CELF-3 (Semel *et al.* 1995) and the PPVT-R (Dunn and Dunn 1981). Language production scores on the CELF-3 (Semel *et al.* 1995). Word Structure and Word Association subtests also showed average or above average performance. Four children did score in the low average

Table 4. Scores for language production, language comprehension, and metaphonology at the follow-up point

	Production[a]		Comprehension[a]		Metaphonology				
Child	Word Structure	Word Association	PPVT-R (1981)	Sentence structure	RP (/15)	AP (/15)	ICD (/15)	FCD (/15)	Total (/60)
Jeanie	37	37	70	50	10	11	5	1	27
Marcy	75	63	22*	37	0	10	3	7	20
Ben	63	91+	52	84	11	1	8	2	22
Faith	91+	98+	94+	75	15	14	13	15	57
Dan	63	16*	78	37	3	14	12	5	34
Lloyd	50	99+	48	50	15	15	15	15	60
Gary	75	25*	78	50	15	15	13	15	58
Brad	37	99+	60	75	13	15	5	4	37
Miles	25*	50	No data	84	11	7	0	0	18
Serena	98+	37	77	75	15	15	11	13	54
Terry	84	75	72	63	15	15	15	15	60
Colin	37	84	45	50	5	14	8	9	36
Mean	61.3	64.5	63.3	60.8	10.7	12.2	9	8.4	40.3
(SD)	(24.01)	(30.6)	(20.2)	(17.3)	(5.3)	(4.3)	(4.9)	(5.9)	(16.7)

[a]Scores: percentile rank scores: *16–25th percentile; +greater than +1 SD.
[b]Word Structure, Word Association and Sentence Structure subtests from CELF-3 (1995).
[c]RP, rhyme production; AP, alliteration production; ICD, initial consonant deletion; FCD, final consonant deletion. Metaphonology is not norm-referenced. Scores are raw scores.

range on one of the tests: Marcy on the PPVT-R, Miles on the Word Structure of the CELF-3, and Gary and Dan on the Word Association subtest of the CELF-3.

In terms of between-task correlations, there were some significant correlations between language and metaphonology scores, i.e. between the metaphonology total score and the PPVT-R (Dunn and Dunn 1981: 0.682, $p=0.021$), and between the metaphonology total score and the CELF-3 Word Structure (Semel *et al.* 1995: 0.739, $p=0.006$). Those two particular language tests also showed significant correlations with each other for these participants (0.717, $p=0.01$).

Verbal memory skills and non-verbal skills (table 5)

In terms of verbal memory, 11/12 children had average or above average scores on the WISC-III (Wechsler 1989) digit span subtest. Eight children had average or above average scores on the Recalling Sentences subtest of the CELF-3 (Semel *et al.* 1995). The digit span and sentence imitation tasks were significantly correlated with one another (0.69, $p=0.013$). Children with scores below the 25th percentile on the digit span subtest were Miles, Marcy (both low average), and Colin (below average). On the Recalling Sentences subtest, children showing below average performance were Colin, Miles, Marcy, and Brad.

There were a number of significant positive correlations between verbal memory scores and other measures. The sentence memory task (CELF-3 Recalling Sentences subtest, Semel *et al.* 1995) and the digit span total score of the WISC-III (Wechsler 1989) both showed significant positive correlations (respectively) with metaphonology total (0.756, $p=0.004$; 0.71, $p=0.01$). There were other positive correlations (1) between the digit span task and the PPVT (Dunn and Dunn 1981) (0.650, $p=0.03$), (2) between the sentence memory task and phonemic discrimination (TOLD-2P, Newcomer and Hammill 1988), and (3) the sentence memory task and the CELF-3 Word Structure subtest (0.677, $p=0.016$).

Table 5. Percentile ranks at follow-up point for verbal memory and non-verbal tasks[a]

Child	Total digit span[b] (WISC-III 1989)	CELF-3 (1995) Recalling sentences	TONI-2 (1990) Geometric categories	TONI-2 (1990) Geometric sequences	TONI-2 (1990) Geometric analogies
Jeanie	37	63	63	37	37
Marcy	16*	5**	50	75	63
Ben	37	75	50	63	37
Faith	98+	95+	63	84	37
Dan	50	37	50	50	50
Lloyd	63	75	50	63	37
Gary	50	50	63	50	75
Brad	37	5**	98+	75	50
Miles	25*	9**	98+	75	50
Serena	50	95+	98+	37	84
Terry	50	50	84	16*	37
Colin	9**	9**	84	37	63
Mean (SD)	43.5 (23.2)	47.3 (34.3)	70.9 (20.2)	55.2 (20.7)	51.7 (16.3)

[a]Scores: 16–25 percentile*; greater than −1 SD**; greater than +1 SD+.
[b]WISC-III (Wechsler 1989) percentile ranks for combined digits forward and digit backward scores.

All children scored within or above normal limits on the three subtests of the TONI-2 (Brown *et al.* 1990): Geometric Categories, Geometric Sequences and Geometric Analogies. One participant had a low average score on Geometric Analogies (Terry).

Academic skills (tables 6 and 9)

On the PIAT-R (Dunn and Marquardt 1989) Reading Recognition (RR) and Reading Comprehension (RC) tasks, 10/12 children had average or above average scores. On the arithmetic subtest, 9/12 children had average or above average performance, and on the spelling subtest, 7/12 children had average or above performance.

Children whose scores fell below the 25th percentile in reading recognition (decoding) were Terry, Brad, Ben, Miles, and Colin; for reading comprehension, Faith, Brad, Ben, Colin, and Miles had scores in that range. For spelling, low and below average scores were noted for Colin, Faith, Dan, Gary, Jeanie, Miles, and Ben.

Follow-up variables and academic skills

A number of follow-up variables were significantly correlated with literacy scores: age, the CELF-3 Word Structure and Recalling Sentences subtests (Semel *et al.* 1995), metaphonology total score, the PPVT-R (Dunn and Dunn 1981), and the Geometric Analogies subtest of the TONI-2 (Brown *et al.* 1990) (tables 7 and 8).

Although the total metaphonology score was correlated with literacy scores, there were only two significant subtask correlations with literacy scores: alliteration production and final consonant deletion. Correlations for alliteration production were 0.662 ($p=0.019$) with decoding, 0.706 ($p=0.01$) with reading comprehension, and 0.669 ($p=0.017$) with spelling. For final consonant deletion, significant correlations were found only with reading tasks: 0.588 ($p=0.044$) with decoding,

Table 6. Percentile ranks for the Peabody Individual Achievement test (Dunn and Marquardt 1989) at the follow-up point

	Grade completed	RR	RC	Spelling	Arithmetic
Jeanie	1	34	27	5**	32
Marcy	Kg	45	42	39	88^{+}
Ben	Kg	21*	21*	2**	73
Faith	1	30	23*	21*	34
Dan	1	34	34	12**	18*
Lloyd	1	34	42	47	34
Gary	1	39	34	8**	79
Brad	1	25*	23*	42	42
Miles	1	12**	8**	5**	21*
Serena	2	95^{+}	94^{+}	>99^{+}	13**
Terry	2	25*	34	34	12**
Colin	2	7**	10**	23*	2**
Mean (SD)		33.4 (22.2)	32.7 (22.2)	27.3 (28.01)	30.7 (24.4)

Percentile by grade for achievement tasks.
RR, reading recognition; RC, reading comprehension. *16–25th percentile; **greater than −1 SD; $^{+}$greater than +1 SD.

Table 7. Significant correlations between follow-up variables and literacy tasks

Variable	Pearson correlations		
	RR	RC	SP
CELF-3 (1995) Word Structure	0.71 (0.001)	0.71 (0.001)	0.60 (0.003)
CELF-3 (1995) Recalling Sentences	0.67 (0.02)	0.67 (0.02)	0.57 (0.053)
Metaphonology (total of all tasks)	0.66 (0.02)	0.72 (0.008)	0.61 (0.035)
Geometric Analogies, TONI-2 (1990)	0.68 (0.01)	0.62 (0.03)	0.66 (0.02)
Age	0.58 (<0.05)	0.64 (0.02)	0.66 (0.02)
PPVT-R (1981)	0.64 (0.03)	0.65 (0.03)	n.s.

PIAT-R (Dunn and Marquardt 1989): RR, reading recognition; RC, reading comprehension; SP, spelling.

Table 8. Stepwise multiple regression concerning the PIAT-R (Dunn and Marquardt 1989) Reading Recognition Subtest

Step	Variable entered	r^2	Change in r^2	Significance of change
1	Word Structure	0.710	0.710	0.001
2	Metaphonology	0.714	0.004	n.s.
3	Age	0.722	0.009	n.s.

and 0.642 ($p=0.024$) with reading comprehension. Neither rhyme production nor initial consonant deletion was significantly correlated with any academic tasks.

In a partial correlation analysis, only one marginally significant correlation remained between concurrent scores: CELF-3 Word Structure and Reading Recognition (significance level of 0.058). Similarly, a multiple regression analysis showed that concurrent metaphonology and age (although strongly correlated individually with literacy tasks), did not contribute significantly to the variance in comparison with CELF-3 Word Structure (table 8).

Pre- and post-intervention scores compared with follow-up study variables

Table 9 shows the few significant correlations between early scores and scores at the follow-up point. Only one preschool study variable was significantly correlated with later reading and spelling skills: post-intervention metaphonology (0.72, $p=0.008$ with total reading, 0.64, $p<0.03$ with spelling). This variable also was also significantly correlated with follow-up metaphonology (0.624, $p<0.03$). The only other significant correlations were between early IPSyn (Scarborough 1990) scores and selected language, verbal memory, and metaphonology scores at the follow-up point. The IPSyn scores were not significantly correlated with literacy scores, however.

Discussion

In the present study, 12 children who had participated in a non-linear phonological intervention study as preschoolers were given a battery of speech, language and

Table 9. Significant Pearson correlation coefficients between preschool and follow-up study variables

	Follow-up variables								
Early variables	MP	CELF-3[a] (1995) RS	CELF-3[a] (1995) SS	PPVT-R (1981)	Digit span[b]	RR[c]	RC[c]	Total R[c]	SP[c]
MP total-post-	0.624[d] (0.03)*					0.73 (0.01)*	0.70 (0.01)*	0.72 (0.008)*	0.64 (0.03)*
IPSyn (1990)[e]-pre-	0.702[g] (0.011)*	0.668 (0.018)*			0.684 (0.014)*				
IPSyn (1990)[e]-post-	0.746[f] (0.005)*		0.631 (0.028)*	0.817 (0.002)*	0.732 (0.007)*				

Pre- and post-intervention phonology PCCs did not show significant correlations with any follow-up variables, nor did pre-metaphonology (MP) scores.
[a]Recalling Sentences (RS) and Sentence Structure (SS) subtests from the CELF-3 (Semel *et al.* 1995).
[b]WISC-III (Wechsler 1989) combined digit forward and digit backwards scores.
[c]Subtests from PIAT-R (Dunn and Marquardt 1989): RR, reading recognition; RC, reading comprehension; total R, total reading score; SP, spelling.
[d]Correlation of total metaphonology scores; however, the primary source of that significant correlation was with follow-up segmentation tasks (0.578, initial consonant deletion. 0.049*; 0.635, final consonant deletion, 0.027*).
[e]IPSyn, Index of Productive Syntax (Scarborough 1990).
[f]Correlation significant with rhyme production only.

academic tasks approximately 3 years later. The purposes of the study were (1) to describe the children's later performance on speech, language and academic tasks as individuals and as a group, and (2) to determine possible relationships between the later performance and variables at the follow-up point, and pre- and post-intervention. Results for the major variables are discussed in turn: phonology and phonemic discrimination, metaphonology, language comprehension and production, non-verbal skills and verbal memory skills, and academic skills. Interactions between variables are discussed within each section, both across and within participants.

Phonology and phonemic discrimination: normalization rate, severity, and interactions

Shriberg *et al.* (1994a, b) suggested age 6 as a cut-off point for short-term normalization in speech development. All children in this study had reached age 6, and 7/12 showed such short-term normalization. The other children were also intelligible in conversation, but showed residual errors on late-acquired phonemes.

Severity has generally been reported to be a fairly reliable predictor of normalization rate (Shriberg *et al.* 1994a, b, Webster *et al.* 1997). In terms of pre-intervention severity, all participants *except* Miles had been predicted to show ongoing impairment in the early school years, because of low PCCs and reductions in word structure. However, only five children (*including* Miles) showed continuing delays in speech development; furthermore, correlations between pre-intervention phonology and phonology at the follow-up were not significant. Thus, pre-intervention severity was not a reliable predictor of ongoing speech impairment in this sample, and could be considered one positive outcome of the earlier study. In terms of post-intervention severity, four of the six children predicted to be at risk for continuing speech impairment did continue to show delayed speech development (Colin, Terry, Marcy and Jeanie). Post-intervention severity was a better predictor of ongoing speech impairment than pre-intervention severity for individuals in this sample, although statistical correlations did not confirm this for the group. In Webster *et al.*'s (1997) similar longitudinal study, about two-thirds of 29 subjects showed normalization, but they had less severe phonological impairments at the outset. Thus, normalization rates for speech impairment were considered favourable in the present study.

Age at onset of intervention was predicted to be relevant for rate of normalization. Colin and Terry were the oldest in the early project (5 at the onset of treatment), and did continue to have delayed speech development at age 8. Jeanie and Marcy, however, were 3 at the onset of treatment, and continued to have (mild) delays at age 6. Thus, age of onset of treatment was not a reliable predictor of normalization rate across the sample.

Considering predisposing factors for slower rates of normalization, four of the children with ongoing speech impairment also had low average scores on the Word Discrimination subtest of the TOLD-2P (Newcomer and Hammill 1988) at the follow-up point (Colin, Terry, Dan, Jeanie). These children may have had a more pervasive phonological impairment that affected both production and perception, even if the effect on perception was minimal. Whether they had compromised underlying representations, or major phonological processing deficits cannot be reliably determined from the data. A speech discrimination in noise task may have given more information. In addition, three children had other speech-related impairments that may have influenced normalization rate. Colin had a mild to moderate fluency disorder at the follow-up point. Both Dan and Terry had observable oral-motor deficits. Dan's speech had a dysarthric quality at the follow-up point, which was not evident in the early intervention period. Terry had lateralized (ungrooved) sibilants. His pronounced overbite, crowded upper teeth, and high palatal arch may have made accurate sibilant production challenging. For Terry, two additional factors may have slowed down normalization: chronic otitis media as a preschooler (with a myringotomy), and a family history of speech impairment (his mother also continued to have lateralized sibilants, and his preschool-aged brother was unintelligible).

Metaphonology: normalization, definition

For the present study, the metaphonology tasks were non-standard and there was no age-based control group. Thus, it is not possible to comment on the 'normalcy' of their performance. There was a general increase in scores with age, however, which suggests the tasks had some chronological validity.

At the follow-up point, two children scored no points on at least one task, and one of those children was one of the older children in the study (Miles). Marcy at age 6;2 was unable to do the rhyme production task and also had a low average score in phonemic discrimination. Her performance on these tasks may have been related, because both require awareness of phonetic similarity. In contrast, Miles at age 7;5 had a superior score on the phonemic discrimination task and a reasonably high score on rhyme production (11/15), yet was unable to do the phonemic segmentation tasks. The results for these two children suggest that phonological awareness may not be a uniform construct, but a compilation of related constructs, which may rely on different subskills. This is perhaps more consistent with Muter and Snowling's (1998) perspective that phonological sensitivity tasks tap independent abilities, than Anthony *et al.*'s (2002) perspective that they reflect one underlying ability (see also the discussion below on academic skills).

In terms of early and later performance, post- (but not pre-) intervention metaphonology was significantly correlated with metaphonology at the follow-up point. The general stability of performance from the post-intervention point to the follow-up point is similar to findings of Webster *et al.*'s (1997) longitudinal study.

The preceding section noted that post-intervention phonology was also predictive of phonology at the follow-up for several children, although the correlation was not statistically significant. The phonology and metaphonology data suggest that *rate of change* during phonological and metaphonological intervention may be predictive of future performance in those domains.

Language comprehension and production: stability across time, and interactions

At the follow-up point, all but three children had scores above the 25th percentile across all language tasks. The PPVT-R (Dunn and Dunn 1981) and the CELF-3 Word Structure subtest were significantly correlated with each other and the total metaphonology score, but not with phonology or other language comprehension tasks.

Significant positive correlations between the early study IPSyn scores (Scarborough 1990) and follow-up language measures were among the few found between early and follow-up scores (table 9). This finding suggests stability of language performance across time for the sample. The significant correlations were somewhat surprising, given the tasks and individual performances. The early IPSyn scores were based on small (perhaps therefore unreliable) language samples and showed delays for 9/12 children, even after intervention. In contrast, the follow-up CELF-3 subtest scores (Semel *et al.* 1995) were in the average or above average range across children. There was a range of performance across children, however, three scoring in the low average range on at least one language test.

Non-verbal skills and verbal memory skills: Homogeneity, diversity, interactions

The TONI-2 (Brown *et al.* 1990) scores (average or above average range) and lack of correlation with language and speech tasks suggest that the group was relatively homogeneous with respect to non-verbal skills involving geometric analogies, categories and sequences.

Verbal memory scores were strikingly more divergent; scores ranged from below to above average. There were a number of significant positive correlations between at least one of the verbal memory scores and other measures, that is with metaphonology total, the PPVT (Dunn and Dunn 1981), phonemic discrimination on the TOLD-2P (Newcomer and Hammill 1988), and the CELF-3 Word Structure subtest. Although there was no group correlation between phonology and verbal memory, two children who had below average scores on at least one verbal memory task at the follow-up point also had below average scores in phonology at that point (Colin, Miles). Overall, phonological working memory did appear to be relevant for performance on a number of verbal tasks.

A cluster of correlated variables emerges from the statistical analyses: verbal memory, metaphonology total, sentence production as tested by the Word Structure subtest of the CELF-3 (Semel *et al.* 1995) (a cloze task), and vocabulary comprehension as tested by the PPVT-R (Dunn and Dunn 1981). This result is congruent with the literature on metaphonology as discussed by other researchers (e.g. Catts *et al.* 1999, Webster *et al.* 1997, Chiappe and Siegel 2000). Metaphonology tasks clearly involve phonological working memory. Cloze tasks also involve working memory, because the carrier phrase must be remembered while words and grammatical morphemes are being retrieved. Vocabulary learning involves short-term memory (at first) and long-term storage. Thus, verbal memory appears to be a core variable for most of these tasks. Similar conclusions were reached in a study of information processing of school-aged children with specific language impairment (Gillam *et al.* 1998). Those authors state that children with specific language impairment 'may have difficulty retaining previously formed phonological codes during multiple mental operations or … [might] avoid creating phonological representations unless such codes are necessitated by task requirements' (Gillam *et al.* 1998: 924). Webster *et al.*'s (1997) study of children with a history of phonological impairment also leads to this interpretation. The present authors would agree with Webster *et al.* that the direction of effect is difficult to prove, however. Is verbal memory development dependent on some other aspect of phonological or language processing, or vice versa? Further large *n* longitudinal studies are needed comparing children with and without language and phonological impairments. A much more well-delineated model of phonological working memory is also needed in order to understand how it works, and what types of tasks might best evaluate its function.

In terms of correlations with early variables, both pre- and post-intervention IPSyn (Scarborough 1990) scores were significantly correlated with digit span at the follow-up point. The import of this is not clear, because the CELF-3 Word Structure subtest (Semel *et al.* 1995) was not significantly correlated with the digit span test at the follow-up point. However, the sentence imitation task was correlated with verbal memory at the follow-up point, and the digit span and sentence imitation tasks were significantly correlated with one another. The small sample size may have negatively affected the correlation data in this instance.

Academic skills

One major purpose of the study was to examine the various interactions between oral speech and language measures and academic performance. The children in the original intervention study were all considered to be at risk for delayed literacy development. Overall, the results were more encouraging than had been anticipated. Ten of the twelve children scored within normal limits on both decoding and comprehension subtests of the PIAT-R (Dunn and Marquardt 1989). On arithmetic tasks, 9/12 had scores within normal limits. Spelling showed lower scores than the other academic tasks; five children had below average scores. This may have suggested that phonological impairments can have more impact on spelling than reading (mirroring findings by e.g. Stackhouse 1982, Clarke-Klein 1994, Bird *et al.* 1995), although there were no significant statistical correlations between any early or concurrent phonology scores and literacy tasks. Thus, there may be other reasons for this result, including a current educational approach in the participant area that does not focus on accurate spelling in the primary grades.

Previous research has shown that severity of phonological impairment, language production, metaphonology, verbal memory, and non-verbal skills, are often correlated with academic performance. In this study, there were no significant correlations between phonology scores and academic skills. Children in the present study had more severe phonological impairments as preschoolers than children in some other studies (Bird *et al.* 1995, Webster *et al.* 1997) yet had more positive results on academic tasks in comparison. Among the non-verbal skills, Geometric Analogies subtest of the TONI-2 (1990) was significantly correlated with all literacy tasks on a zero order correlation, showing partial congruence with Bird *et al.*'s (1995) study, which showed positive correlations between non-verbal intelligence and literacy tasks. However, this variable became non-significant in a partial correlation analysis, possibly because the group performed generally well on non-verbal tasks.

Both post-intervention and concurrent metaphonology skills were highly correlated with academic skills. If children did not benefit sufficiently from early metaphonological intervention, they were more likely to be delayed in literacy development. These results replicate those of other researchers, e.g., Magnusson and Naucler (1989, 1993, 1998), Bird *et al.* (1995), Webster *et al.* (1997), and Gillon (2002). Not all metaphonology tasks were significantly correlated with reading, however, rather, only alliteration production and final consonant deletion. This divergence in literacy scores and correlations supports the notion that metaphonological skills may tap different underlying abilities as suggested by Muter and Snowling (1998). Furthermore, only the post-treatment metaphonology score remained highly correlated with reading scores. The concurrent score contributed less to the overall variance, possibly because the children had had more exposure to phonological awareness tasks by the follow-up testing point through schooling. In both ordered correlations and multiple regression analyses, language production as tested by the CELF-3 (Semel *et al.* 1995) Word Structure subtest showed the strongest correlations with one literacy task (decoding). This finding is congruent with research by Catts *et al.* (1999) and others that shows language production to be highly correlated with literacy development across children.

Individual performance

There was no one-to-one correspondence between performance on academic and other tasks across individuals, consistent with findings in the literature. The small sample size, and different ages and literacy experiences of the children undoubtedly contributed to that variability. A number of individuals stand out in the sample.

In terms of positive outcomes, Lloyd scored within normal limits or above on all tasks. As one of the youngest participants in the earlier study, his performance was exemplary of short-term normalization across domains. Serena had high scores on literacy tasks (over two standard deviations above the norm), but continued to have residual phonological errors (primarily /r/) and below average scores on arithmetic. Both Serena and Lloyd had superior scores on at least one verbal memory task, exemplifying the link between verbal memory and literacy outcomes, as found in the group correlations. Serena also scored at the highest percentile rank on the CELF-3 Word Structure subtest (Semel *et al.* 1995), another strongly correlated variable with literacy for the group.

Colin and Miles had below average performance or low average performance on all academic tasks, and each had many related factors, both at the follow-up point and in the early study. At the follow-up point, both had below average verbal memory scores and ongoing phonological impairments. In addition, Miles had a low average score on the *CELF*-3 Word Structure subtest (Semel *et al.* 1995), and the lowest metaphonology score of the group. He could not do phonemic segmentation tasks, which are considered by some researchers to be the best predictors of reading ability (Hulme *et al.* 2002). Comparing Miles's performance with that of another child, Marcy, this perspective is supported in this small sample. Marcy could do phonemic segmentation tasks, but could not do rhyming production tasks, and she performed within normal limits on all literacy tasks. This individual divergence between literacy scores and metaphonology tasks suggests that metaphonological tasks tap different abilities as is also suggested by the group data.

In the early study, Colin had the most severe phonological impairment in the early study (a PCC of 10.9% at age 4;11). The severity of his phonological impairment, his age at onset of treatment, and his delay in language production predicted ongoing impairment in speech and literacy, both of which were confirmed in the follow-up study. Miles, however, had only a moderate delay in phonological development at age 4, and average IPSyn (Scarborough 1990) scores in language production after intervention. These factors did not predict later impairments in speech and literacy. However, he had shown minimal gain in metaphonology across the early study; thus, his literacy outcomes were congruent with the general significant correlations between post-intervention metaphonology and literacy development, as predicted.

Conclusion

Overall, the results of this study were more positive than anticipated for this admittedly small cohort of children. Most children performed within normal limits on a number of speech, language and literacy tasks, in spite of their early history of severe phonological impairments and delays in language production and metaphonology. Whether this positive outcome will hold up in later school years is difficult to predict, given the increasing complexity of interactions between child, family, and educational variables. However, a similar cohort of six children from a previous non-linear intervention study (Bernhardt 1990) have graduated from high school as literate and intelligible young adults, even though three of the six had difficulties with literacy tasks throughout their schooling that meant they had to work hard to accomplish this goal. This suggests that the current cohort might also overall continue to have positive outcomes in speech, language and literacy.

Previous research has identified verbal memory, phonemic discrimination, phonology, language production, metaphonology and non-verbal skills as potential correlates with academic skill development. Only two children had across-the-board delays in academic skill acquisition in this study, and they each had below average performance on some of the identifiable predictors: verbal memory, residual phonological impairments, metaphonology, and/or language production. Replicating Catts *et al.* (1999) and others, language production was the most highly correlated concurrent variable with reading decoding in this study. Post-intervention metaphonology was the only predictive variable for literacy skills from the preschool intervention study across participants, replicating many studies in the area of phonological awareness.

In this study, the early intervention focused on the multiple levels of the phonological hierarchy, and provided direct metaphonological instruction. These approaches both emphasize the structure of words, and in so doing, may have facilitated both phonological and literacy development. Gillon (2002) reached the same conclusion about her study of phonological awareness intervention as noted in the introduction to this paper. In terms of clinical implications, intervention outcomes studies which have examined the effects of metaphonological and phonological intervention (e.g. Major and Bernhardt 1998, Gillon 2002, Bernhardt and Major, current) suggest that explicit metaphonological instruction may be necessary for at least some children to acquire literacy skills if they have history of phonological impairment. However, there is considerable individual difference among children in their ability to gain metaphonological awareness and subsequent literacy skills without explicit instruction in metaphonology during phonological intervention (Major and Bernhardt 1998, Hesketh *et al.* 2000, Bernhardt and Major, present study). This implies that it may not *always* be necessary to provide direct training for metaphonological skills during phonological intervention, but rather to evaluate a child's metaphonological skills as part of phonological assessment. If a child does not appear to show gains in metaphonology as a result of phonological intervention, direct instruction in metaphonology may be necessary in order to promote access to the written code. Early phonological and metaphonological intervention can promote normalization of speech development, and normal acquisition of literacy skills for children with severe phonological impairments. Not all children will demonstrate short-term normalization, but the level and prevalence of risk can be reduced.

Acknowledgements

The authors acknowledge the children and their families for their participation in the follow-up study. The authors also thank the UBC's Humanities and Social Sciences Small Grants for funding support, and the referees of the present paper for helpful comments.

Appendix: Overview of the earlier non-linear phonological intervention study

In the non-linear phonological intervention study reported by Major and Bernhardt (1998), 19 preschool children participated along with their significant others in 45 min individual treatment sessions three times a week over 16 weeks with community speech–language pathologists (SLPs). The first 12 weeks addressed phonological goals, divided equally between syllable structure and segments (phonemes) and features. The final 4 weeks also addressed metaphonological goals in rhyming, alliteration and/or segmentation.

An alternating treatment design and a 'cycles' approach was used to address goals derived from the non-linear phonological analyses for (1) syllable and word structure and (2) features and their related segments (phonemes). (For detailed examples of analyses, see Bernhardt 1990, 1992, 1994b, Bernhardt and Gilbert 1992, and Bernhardt and Stemberger 2000). Goals for each child were based on individual needs as determined by the analyses, in accordance with three basic principles. The first principle was to use strengths in the multilevel phonological system to address needs. When addressing syllable structure goals, only segments (phones) from the pretreatment inventory were used. Similarly, new features and segments were targeted only in the child's existing syllable and word structures. When a child was not immediately stimulable for a structure or segment, shaping from an existing structure or segment was used to approximate the new target.

The second principle was to address phonological categories (features, structures) rather than individual segments. Thus, in targeting the category of fricatives (feature [+continuant]), two or three fricatives contrasting in place and voicing would be included in treatment. When addressing the category of CCVC, a number of clusters would be included (e.g. /tw/, /kw/, /bj/).

The third principle was to target non-default (marked infrequent) structures and features rather than default (unmarked frequent) structures and features as much as possible (i.e. /ʃ/ rather than /s/, /v/ rather than /f/, /w/-clusters rather than /s/-clusters). This principle was relaxed, however, where broad developmental considerations appeared more logical (e.g. CVC before CCVC if a child had neither; alveolars over dentals if a child had neither), or if a child appeared to need early success in the treatment process, as might be expected for a developmentally earlier target.

Treatment approaches included both phonological awareness and imitation activities with an emphasis on stories and active games tailored to the individual child. SLPs received 2 days of training to learn the analysis and treatment procedures (which included creation of new activities). Treatment techniques designed to address syllable and word structure focused on subsyllabic components: onsets (consonants before the vowel), rhymes (vowel and following consonants) and syllable 'weight units' or 'moras' (see Bernhardt 1994a, and Bernhardt and Stemberger 2000 for examples of activities, also Major and Bernhardt 1998, appendix, for more general information on the treatment process).

In the last 4-week block of the intervention study (12 sessions), there was also direct instruction in metaphonological awareness, with activities focusing on rhyming and alliteration in particular.

References

Anthony, J., Lonigan, C., Burgess, S., Driscoll, K., Phillips, B. and Cantor, B., 2002, Structure of preschool phonological sensitivity: overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology*, **82**, 65–92.

Bernhardt, B., 1990, Application of nonlinear phonological theory to intervention with six phonologically disordered children. Unpublished doctoral dissertation, University of British Columbia, Vancouver.

Bernhardt, B., 1992, The application of nonlinear phonological theory to intervention with one phonologically disordered child. *Clinical Linguistics and Phonetics*, **6**, 283–316.

Bernhardt, B., 1994a, Phonological intervention techniques for syllable and word structure development. *Clinics in Communication Disorders*, 54–65.

Bernhardt, B., 1994b, The prosodic tier and phonological disorders. In M. Yavas (ed.), *First and Second Language Phonology* (San Diego: Singular), pp. 149–172.

Bernhardt, B., Edwards, S. and Rempel, L., 1995, Metaphonological skills of typically developing 3-year-olds and children with phonological disorders. In T. Powell (ed.), *Pathologies of Speech and Language: Contributions of Clinical Phonetics and Linguistics* (New Orleans: ICPLA), pp. 135–142.

Bernhardt, B. and Gilbert, J., 1992, Applying linguistic theory to speech–language pathology: the case for nonlinear phonology. *Clinical Linguistics and Phonetics*, **6**, 123–145.

Bernhardt, B. H. and Stemberger, J. P., 2000, *Workbook in Nonlinear Phonology for Clinical Application* (Austin: Pro-Ed).

Bernhardt, B. H. and Stemberger, J. P., 2001, Intervocalic consonants in the speech of English-speaking Canadian children with phonological disorders. *Clinical Linguistics and Phonetics*, **16**, 199–214.

Bernthal, J. and Bankson,, 1998, *Articulation and Phonological Disorders*, 4th edn (Englewood Cliffs: Prentice-Hall).

Bird, J., Bishop, D. V. M. and Freeman, N. H., 1995, Phonological awareness and literacy development in children with expressive phonological impairments. *Journal of Speech and Hearing Research*, **38**, 446–462.

Bishop, D. V. M. and Adams, C., 1990, A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, **31**, 1027–1050.

Bopp, K., 1995, The effects of phonological intervention on morphosyntactic development in preschool children with phonological and morphosyntactic disorders. Unpublished Master's thesis, University of British Columbia, Vancouver.

Brady, S., Shankweiler, D. and Mann, V., 1983, Speech perception and memory coding in relation to reading ability. *Journal of Experimental Psychology*, **34**, 345–367.

Brown, L., Sherbenou, R. and Johnsen, S., 1990, *Test of Nonverbal Intelligence—2* (Austin: Pro-Ed).

Bryant, P., 2002, It doesn't matter whether onset and rime predicts reading better than phoneme awareness does or vice versa. *Journal Of Experimental Child Psychology*, **82**, 41–46.

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

Catts, H., 1993, The relationship between speech–language impairments and reading disabilities. *Journal of Speech and Hearing Research*, **36**, 948–958.

Catts, H., Fey, M., Zhang, X. and Tomblin, J. B., 1999, Language basis of reading and reading disabilities: evidence from a longitudinal investigation. *Scientific Studies of Reading*, **8**, 331–361.

Chiappe, P. and Siegel, L. S., 2000, Working memory, inhibitory control, and reading disability. *Memory and Cognition*, **78**, 8–17.

Clarke-Klein, S., 1994, Expressive phonological deficiencies: impact on spelling development. *Topics in Language Disorders*, **14**, 40–55.

Dodd, B., 1995, *Differential Diagnosis and Treatment of Children with Speech Disorder* (London: Whurr).

Dunn, L. and Dunn, L., 1981, *Peabody Picture Vocabulary Test—Revised* (Circle Pines: American Guidance Service).

Dunn, L. and Marquardt, F., 1988, *The Peabody Individual Achievement Test—Revised* (Circle Pines: American Guidance Service).

Fazio, B., 1986, Serial memory in children with specific language impairment. *Topics in Language Disorders*, **17**, 58–71.

Gathercole, S. E. and Baddeley, A. D., 1993, *Working Memory and Language* (Hove: Lawrence Erlbaum).

Gierut, J., 1998, Treatment efficacy: functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research*, **41**, S85–S100.

Gillam, R., Cowan, N. and Marler, J., 1998, Information processing by school-age children with specific language impairment: evidence from a modality effect paradigm. *Journal of Speech, Language, and Hearing Research*, **41**, 913–926.

Gillam, R. and Van Kleeck, A., 1996, Phonological awareness training and short-term working memory: clinical implications. *Topics in Language Disorders*, **17**, 72–81.

Gillon, G., 2002, Follow-up study investigating the benefits of phonological awareness intervention for children with spoken language impairment. *International Journal of Language and Communication Disorders*, **37**, 381–400.

Gillon, G. and Dodd, B., 1994, A prospective study of the relationship between phonological, semantic and syntactic skills and specific reading disability. *Reading and Writing*, **6**, 321–345.

Gillon, G. and Dodd, B., 1995, The effects of training phonological, semantic, and syntactic processing skills in spoken language of reading ability. *Language, Speech and Hearing Services in Schools*, **26**, 58–68.

Goldman, E. and Fristoe, M., 1986, *Goldman Fristoe Test of Articulation—Revised* (Circle Pines: American Guidance Service).

Goswami, U., 2002, In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams, and Stuart. *Journal of Experimental Child Psychology*, **82**, 47–57.

Gottardo, A., Stanovich, K. E. and Siegel, L. S., 1996, The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology*, **63**, 563–582.

Green, M., 1991, Phonological awareness skills in phonologically disordered and phonologically normal preschoolers. Unpublished Master's thesis, Syracuse University, New York.

Hesketh, A., Adams, C., Nightingale, C. and Hall, R., 2000, Phonological awareness therapy and articulatory training approaches for children with phonological disorders: a comparative outcomes study. *International Journal of Language and Communication Disorders*, **35**, 337–354.

Hodson, B., 1986, *Assessment of Phonological Processes—Revised* (Austin: Pro-Ed).

Hulme, C., Hatcher, P., Nation, K., Brown, A., Adams, J. and Stuart, G., 2002, Phoneme awareness is a better predictor of early reading skill than onset–rime awareness. *Journal of Experimental Child Psychology*, **82**, 2–28.

Khan, L. and Lewis, N., 1986, *Khan–Lewis Phonological Analysis* (Circle Pines: American Guidance Service).

Lewis, B. and Freebairn, L., 1992, Residual effects of preschool phonology disorders in grade school, adolescence and adulthood. *Journal of Speech and Hearing Research*, **35**, 819–831.

Magnusson, E. and Naucler, K., 1989, The development of language and linguistic awareness in language disordered and normally speaking children as related to reading and spelling acquisition. *Praktisk Lingvistik*, **13**, 13–30.

Magnusson, E. and Naucler, K., 1993, The development of linguistic awareness in language-disordered children. *First Language, Special Issue: Language Development in Special Populations*, **13**, 93–111.

Magnusson, E. and Naucler, K., 1998, Reading and writing development: report from an ongoing longitudinal study of language-disordered and normal groups from pre-school to adolescence. *Folia phoniatrica et Logopaedica*, **50**, 271–282.

Major, E. M. and Bernhardt, B., 1998, Metaphonological skills of children with phonological disorders before and after phonological and metaphonological intervention. *International Journal of Language and Communication Disorders*, **33**, 413–444.

McBride-Chang, C., 1995, What is phonological awareness? *Journal of Educational Psychology*, **87**, 179–192.

Muter, V. and Snowling, M., 1998, Concurrent and longitudinal predictors of reading: the role of metalinguistic and short-term memory skills. *Reading Research Quarterly*, **33**, 320–337.

Naucler, K. and Magnusson, E., 1999, Reading and writing development. A longitudinal study from preschool to adolescence: status report. *Working Papers of the Department of Linguistics, Lund University*, **47**, 169–180.

Newcomer, P. and Hammill, D., 1988, *Test of Language Development—2 Primary* (Austin: Pro-Ed).

Olswang, L., 1998, Treatment efficacy research. In C. Frattali (ed.), *Measuring Outcomes in Speech Language Pathology* (New York: Thieme), pp. 134–150.

Reynell, J., 1985, *Reynell Developmental Language Scales—Revised* (Windsor: NFER-Nelson).

Scarborough, H., 1990, Index of productive syntax. *Applied Psycholinguistics*, **11**, 1–22.

Semel, E., Wiig, E. and Secord, W., 1995, *The Clinical Evaluation of Language Fundamentals—3* (San Antonio: Psychological Corporation).

Shriberg, L., Gruber, F. A. and Kwiatkowski, J., 1994a, Developmental phonological disorders III: Long-term speech–sound normalization. *Journal of Speech and Hearing Research*, **37**, 1151–1177.

Shriberg, L. and Kwiatkowski, J., 1982, Phonological disorders III: a procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders*, **47**, 256–270.

Shriberg, L. and Kwiatkowski, J., 1988, A follow-up study of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders*, **53**, 144–156.

Shriberg, L. and Kwiatkowski, J., 1994, Developmental phonological disorders I: A clinical profile. *Journal of Speech and Hearing Disorders*, **51**, 140–161.

Shriberg, L., Kwiatkowski, J. and Gruber, F. A., 1994b, Developmental phonological disorders II: Short-term speech–sound normalization. *Journal of Speech and Hearing Research*, **37**, 1127–1150.

Snow, C., Burns, M. and Griffin, P., 1998, *Preventing Reading Difficulties in Young Children* (Washington, DC: National Academy Press).

Snowling, M., van Wagtandonk, B. and Stafford, C., 1988, Object naming and deficits in dyslexia. *Journal of Research in Reading*, **11**, 67–82.

SPSS, 2001, *SPSS Graduate Version 11.0 for Windows* (Englewood Cliffs: Prentice-Hall).

Stackhouse, J., 1982, An investigation of reading and spelling performance in speech disordered children. *British Journal of Disorders of Communication*, **17**, 53–60.

Stoel-Gammon, and Dunn, C., 1985, *Normal and Disordered Phonology in Children* (Austin: Pro-Ed).

Webster, P. E. and Plante, A. S., 1992, Effects of phonological impairment on word, syllable, and phoneme segmentation and reading. *Language, Speech, and Hearing Services in Schools*, **23**, 176–182.

Webster, P. E., Plante, A. S. and Couvillian, L., 1997, Phonologic impairment and pre-reading: update on a longitudinal study. *Journal of Learning Disabilities*, **30**, 365–375.

Wechsler, D., 1989, *Wechsler Preschool and Primary Intelligence Scale—Revised* (New York: Psychological Corporation).

Whitehurst, G. and Lonigan, C., 1998, Child development and emergent literacy. *Child Development*, **69**, 848–872.

Zimmerman, I., Steiner, V. and Pond, R., 1992, *The Preschool Language Scale—3* (Columbus: Merrill).