Characterizing the Growth Trajectories of Language-Impaired Children Between 7 and 11 Years of Age

REPORT

James Law
Queen Margaret University, Edinburgh

J. Bruce Tomblin
Xuyang Zhang
University of Iowa

Background: A number of different systems have been suggested for classifying language impairment in children but, to date, no one system has been widely accepted.
Method: This paper outlines an alternative system looking for distinct patterns of change in receptive language skills across time, involving a secondary analysis of children identified as having specific language impairment.
Participants: The participants were 184 children age-assessed at 3 time points—7, 8, and 11 years of age.
Results: The pattern of receptive language development is highly predictable. The dominant pattern of growth is consistent with declining rates of growth over time for all children. The primary way in which the children differ is with respect to their initial severity. The testing of the 2 classification systems revealed some statistically significant differences among the subtypes with regard to the shape of the growth rates, but the effect sizes associated with these differences were very small. Thus, it is possible to conclude that beyond the dominant pattern of growth, some subtypes of language impairment at 7 years of age showed only subtle differences in receptive language change across time. The results are discussed in terms of the sample selection and the age of the children who were studied.

KEY WORDS: growth trajectories, language impairment, children

A common problem within both clinical and research work on communication disorders concerns the classification of individuals into subtypes of communication disorder. This interest in classification, or *taxonomy,* can be found in all areas of science, whether it was the development of the periodic table initially constructed by Mendeleyev or the Linnean system for classifying life forms. In all cases, certain properties of the natural entities of interest are used in the taxonomic system to serve as the basis of the organizational system and become important aspects of the theoretical account. Ultimately, the nature and organization of these categories, or *taxa,* according to these properties should reveal important aspects of the phenomena to be explained.

There have been a number of attempts to generate classification systems for developmental language impairment (Aram, Ekelman, & Nation, 1984; Bishop & Rosenblum, 1987; Rapin, 1996; Wilson & Risucci, 1986; World Health Organization, 1992), but no one system is commonly used. Most recently, van der Lely and colleagues (van der Lely, 2003; van der Lely, Rosen, & McClelland, 1998) proposed a subtype of primary language impairment involving just a grammatical deficit, although others have been unable to validate this (Bishop, Bright, James, Bishop, & van der Lely, 2000). The only available replication of one of these taxonomies that

we were able to detect involved the system developed by Rapin and Allen (1987; Rapin, 1996). Conti-Ramsden and colleagues used cluster analysis to identify six groups of children with language impairment and mapped five of them onto Rapin and Allen's (1987) classification system (Conti-Ramsden, Crutchley, & Botting, 1997). These systems all share the assumption that the subtypes of primary language impairments are qualitatively distinct in their pattern of performance. In contrast, Bishop and colleagues (Bishop, 1994; Bishop et al., 2000) have suggested that although there may be a core inherited disorder, there may not be a set of distinct diagnostic entities but, rather, these may be arrayed in a multidimensional space. In concert with this viewpoint, it has been argued—using data from a population-based rather than a clinical sample—that rather than strict subtypes, developmental language disorder may be best described as regions within at least a three-dimensional space containing individual differences with respect to language abilities, including those that comprise language disorder (Tomblin, Zhang, Catts, Ellis Weismer, & Weiss, 2004). Two classes of children were identified across language-impaired and typically developing children. The first of these classes had relatively poor semantic/syntactic skills and relatively poor phonological processing/verbal short-term memory but relatively good social communication skills. The second class comprised children with relatively weak social communication skills. It has been suggested that there is probably a single dimension accounting for language with test-specific or language area–specific minor dimensions at different ages (Tomblin & Zhang, 2006). Of particular significance was what appears to be a change across time with the single construct giving way to a two-dimensional model, with distinct vocabulary/sentence use dimensions emerging in the later years. Recently, Dollaghan (2004) directly tested whether specific language impairment (SLI) represented a categorically different form of language variation from normal forms or whether it was a region along a dimension. In this study, a taxometric analysis was performed on the language scores of 3- and 4-year-old children. In both groups, there was no evidence in support of SLI being a discrete variant of primary language impairment but, rather, SLI appeared to fall within the same dimensions as children with normal levels of language.

One of the key challenges to the development of classification systems for developmental language impairment has been their capacity to capture a stable subtype or trait across time—effectively, their predictive validity. Conti-Ramsden and Botting (1999) evaluated the stability of their Rapin and Allen (1987) subtypes after 1 year. Although they found a similar set of subtypes among their participants, there was considerable instability of membership in these subtypes over time. Stothard,

Snowling, Bishop, Chipchase, and Kaplan (1998) found considerable variability in class membership between SLI and general delay in children who were followed from 5.5 to 15–16 years of age.

Part of the value of a clinical taxonomic system is its potential to reveal associated conditions that have different but systematic relationships with the subtypes. Such evidence provides more support for the distinctiveness of the subtypes at both a phenomenological and etiological level. There have now been a series of well-documented studies that have tracked language over time and that have examined outcomes associated with developmental language impairment. The studies to which reference is most commonly made are those in Dunedin, New Zealand (Silva, Williams, & McGee, 1987); Walthamstow, United Kingdom (Richman, Stevenson, & Graham, 1982); Ottawa, Carleton, Canada (Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996; Johnson et al., 1999) and Iowa, United States (Tomblin et al., 1997). Taken together, these studies have demonstrated high levels of co-morbidity among language, literacy, and behavior in clinical populations, such that early language learning difficulties should best be regarded as a risk factor for subsequent educational underattainment and that there is also a high level of continuity across both the primary school and the teenage years. This literature has also shown that outcomes vary with respect to certain subtypes of early speech and language impairment. Beitchman and colleagues (1996) have shown that both behavioral and academic outcomes are poorer for children with language problems and, more specifically, children with receptive language problems than for children with speech sound disorders only. Botting and Conti-Ramsden (2000) found behavioral outcomes to be differentiated according to subtypes based on modality such that children with receptive–expressive impairments did more poorly than those with impairments restricted to expressive language. These differential outcomes could be used as evidence that the subtypes identified in these studies represent different subtypes of language or that poor outcomes are simply a function of severity.

Growth Trajectories and Taxonomies

In recent years, researchers have begun to characterize language development and developmental language impairment across multiple time points. These methods allow researchers to characterize the dynamic growth pattern or trajectory as a characteristic of the language impairment in its own right. Whether these can truly be termed *growth curves* depends on the number of data points available for analysis. The notion that these developmental trajectories are important features in language impairment can be found in Karmiloff-Smith's

(1998) admonition that the key to understanding developmental disorders is in understanding development—that is, growth and change. Leonard (1998) exemplified this kind of thinking when he suggested that there may be a number of potential pathways for children identified with language impairments. Normal language development is characterized as a linear function of time, although there are some suggestions that, at least for some grammatical functions, performance levels out as children achieve competence. Three possible patterns are proposed. The first possibility is that children with language impairment may also be characterized as starting at the same point as the typically developing children, with their scores diverging down across time relative to children with typical development. This could be characterized as the *deterioration hypothesis*. A second possibility is that these children develop at the same speed as typically developing children but then plateau at a specific point, not developing any further. This could be characterized as the *plateau hypothesis*. Both of these two possibilities would effectively represent a decrease in the children's language scores across time. A third possibility, which might be termed the *tracking hypothesis*, not considered by Leonard, would be that language-impaired children are simply slower to take off, but once started, their language development would be parallel to that of typically developing children. They do not get better or worse but tend to stay on the same trajectory. Rice and colleagues (Rice, Wexler, & Hershberger, 1998) have provided data showing this kind of growth pattern with respect to the growth of specific grammatical forms and vocabulary in children with and without SLI. Similarly, Zhang and Tomblin (2004) have found little heterogeneity in the growth of language in a large number of children who varied considerably with respect to language abilities at school entry. These findings would suggest that there is very little heterogeneity in the growth characteristics of children with language impairment and, furthermore, the growth in children with language impairment, at least in the school years, is quite similar to that of typically developing children.

The aim of the present study was to look at the patterns of language change in a group of children with significant clinical levels of language disorder who participated in the original Conti-Ramsden Manchester Language Study (Conti-Ramsden & Botting, 1999). The emphasis in this article is on the children's receptive language skills rather than their expressive language skills or their speech skills. Receptive language measures were used in this study because they were available for all children, at all time points examined, whereas expressive measures were available only for a subset of children at each time point. Two analyses were performed in order to address two questions: (a) Are the growth trajectories of these children heterogeneous? and (b) Do the growth trajectories of subgroups of these children differ? To address the first question, the growth characteristics of each member of the group were characterized using a single language measure across time. The measure in question was The Test of Reception of Grammar (TROG; Bishop, 1982), a measure of comprehension that is widely used in the clinical research literature. To address the second question, the trajectories of subgroups of these children were compared using the same measure. Two systems of subcategorization were used. The first system corresponded to a traditional expressive/receptive language impairment classification system of the type adopted in the ICD-10 classification system (World Health Organization [WHO], 1992), and the second system corresponded to the Rapin and Allen (1987) classification system outlined previously (see Table 1). In both cases, the groups were operationalized in the first instance by the authors of the original Conti-Ramsden and Botting study (Conti-Ramsden & Botting, 1999).

Method

The study was conceptualized as a secondary data analysis of a relatively large cohort of children identified as having primary language impairment. The cohort in question, from the Conti-Ramsden Manchester Language Study, is the single largest clinically identified cohort of children with pronounced speech and language impairments.

Table 1. Group characteristics of Conti-Ramsden et al. (1997) clusters and Rapin and Allen (1987) subtypes.

Conti-Ramsden Cluster	Description	Rapin and Allen equivalent
1	Good articulation but other language skills are poor.	Lexical–syntactic deficit syndrome
2	Poor word reading.	None
3	General poor language but with good naming vocabulary.	Verbal dyspraxia
4	Poor articulation and phonology. Normal language.	Phonological programming deficit syndrome
5	Articulation is fair, but performance on all other tests is poor.	Phonological syntactic deficit syndrome
6	Good articulation and reasonable expressive language but poor receptive language.	Semantic–pragmatic deficit syndrome

Participants

The data set comprises longitudinal information on a subset of a cohort of 242 children with SLI. At age 7 years, these children were sampled from the population attending language units in England. These children were reassessed at 8 and 11 years of age. The children have been followed through to 14 years of age, but it is the earlier phase of their development that is the focus of this analysis. This cohort has already been extensively discussed in the literature in terms of its initial characteristics (Conti-Ramsden & Botting, 1999), their outcomes at 11 years of age (Conti-Ramsden, Botting, Simkin, & Knox, 2001), the extent to which it is possible to predict outcomes (Botting, Faragher, Simkin, Knox, & Conti-Ramsden, 2001; Conti-Ramsden & Botting, 1999; Conti-Ramsden, Botting, & Faragher, 2001), and subgroup children across time (Conti-Ramsden, Crutchley, & Botting, 1997). This article only reports results from 196 participants who had a TROG score on all three time points, and for some analyses, the focus will be on those 184 participants for whom there were sufficient data to apply the operationalization of the ICD-10 and both the Rapin and Allen (1987) classification and Conti-Ramsden clusters.

Table 2 provides descriptive statistics concerning the TROG scores. These data show that the composite measure systematically increases across the age intervals and has a similar variance at each age interval. The tests used to operationalize the different categories of speech and language impairment were the TROG, three subtests of the British Ability Scales (Number Skills, Naming Vocabulary, and Word Reading; Elliott, 1983), the Goldman–Fristoe Test of Articulation (Goldman & Fristoe, 1986), the Bus Story (Renfrew, 1991), and Raven's Coloured Progressive Matrices (Raven, 1986). The first method of subclassification adopted by Conti-Ramsden and Botting (1999) might be characterized as comparable to the ICD-10 system. In this case, four categories were identified—namely, those with normal language scores (above the 40th percentile on the language measures), those with expressive difficulties only, those with expressive/receptive difficulties combined, and a category of children with "complex/receptive" difficulties. The second, third, and fourth categories were comparable to, although not strictly derived from, the language impairment categories identified in the ICD-10. Group 1 was made up of children with language scores within the normal range. This included children who had been identified as having significant difficulties at Time 1 but who had then been transferred to mainstream school and whose difficulties had not been picked up on standardized language measures at Time 2 and Time 3. Group 2 had comprehension within the normal range but poor expressive language skills. Group 3 had both expressive and receptive language scores outside the normal range. Group 4 had poor expressive language scores but lower receptive language scores (i.e., below the 10th percentile). This last group was designated as having complex SLI because the children appeared to have relatively strong expressive language skills.

In the second method of subclassification, the procedure followed was the same as that described in Conti-Ramsden et al. (1999—that is, the data set was analyzed using cluster analysis and the results indicated that, in part at least, the resulting categories reflected those identified by Rapin, 1996). The starting point is the three clinical groupings described in the preceding paragraph, each of which is then broken down into two groups. Thus, the expressive language disorder group is further broken down into verbal dyspraxia and phonological programming difficulties. The expressive SLI grouping is broken down into verbal auditory agnosia and phonological programming deficit syndrome. The complex SLI group (or those in Rapin's "higher order processing disorder") is broken down into lexical–syntactic deficit syndrome and semantic–pragmatic deficit syndrome. The Conti-Ramsden Clusters 1–6 correspond to the six Rapin groupings in the original analysis, although Conti-Ramsden and colleagues identified a group whose difficulties appeared to have resolved (language scores above the 40th percentile), and Rapin and Allen (1987) identified a group with verbal auditory agnosia, which was not detected in the Conti-Ramsden sample. In the original analysis, psychometric data were supported by teacher report. The six clusters are provided in Table 1.
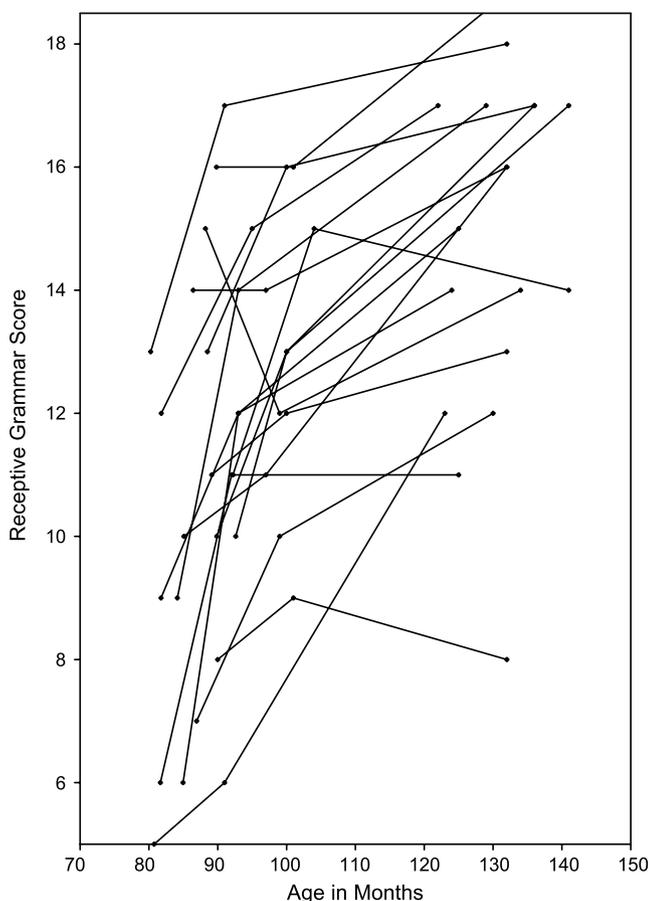
Results

The first question to be addressed concerned the language growth characteristics of the group as a whole and the extent to which there were differences in the magnitude of the growth among the children in this sample. Table 2 provides the descriptive statistics for the TROG scores for the total group at each of the three observation intervals. As expected, the central tendencies reflected

Table 2. Means, medians, and standard deviations of the Receptive Language score (The Test of Reception of Grammar [TROG]; Bishop, 1982) at 7, 8, and 11 years of age.

Variable	Age		
	7 years	8 years	11 years
M	10.10	12.18	15.16
Mdn	10	12	16
SD	3.46	3.21	2.99
n	196	196	196

by the means and medians increased across the longitudinal study, and the variance in TROG reduced somewhat with age. The linear increase in the scores was found to be significant, $F(1, 367) = 587.39, p < .0001$, when these data were analyzed using mixed linear model within SAS.

The extent to which there were individual differences can be seen, however, in Figure 1. This figure contains the individual growth functions for every 10th individual when the participants were ordered from the lowest initial score to the highest. Therefore, these 18 individuals represent a stratified sampling of individuals according to the severity of their language problem at age 7 years. These individual growth patterns reveal a dominant trend for the language skills of all individuals to increase. There are some variations in the manner of this growth that result in some trends being linear, whereas others appear to have an accelerating exponential or decelerating logarithmic shape. The extent to

which these shapes are reliable, however, cannot be determined, as they may simply reflect measurement error at different time points. At least four time points would be needed to detect nonlinear patterns.

Receptive Language Change for the Expressive/Receptive Subtypes of the ICD-10 Classification System

Although the data did not allow us to test the extent to which individual patterns of language change differed, the data did allow for tests of differential growth rate among predefined groups of children. The mean test performance for each group and at each data point are provided in Table 3.

In each case, the children followed the predicted incremental increase in language skill across time. The growth trajectory for each group is then plotted in Figure 2. Groups 1–3 always followed the same pattern, with Group 1 performing better than Groups 2 and 3. Despite the lower receptive language scores at age 7 years, Group 4 functioned close to Group 2 at 8 and 11 years of age.

To test whether these groups had similar growth characteristics, the language scores were submitted to a mixed model within SAS. This analysis method provides a means to perform multilevel analysis. When the data are longitudinal such as the data in this study, this analysis can be considered a *growth curve analysis,* in which age was a repeated measure (random effect) and group was a between-subjects measure (fixed effect). Because only three time points were available, only a linear model with an intercept and linear slope was used. This analysis showed that across all groups, there was a significant effect of age, $F(1, 364) = 382.72, p < .0001, \text{T}^2 = 0.16$. There

Figure 1. Receptive language scores across three observation intervals for every 10th child sampled systematically from the lowest initial score to the highest initial language level, thus providing a representation of the range of initial language ability.

Table 3. Means and standard deviations of the Receptive Language score (TROG) for the four language subgroups comparable to those in ICD-10 at 7, 8, and 11 years based on clusters identified in Conti-Ramsden, Crutchley, & Botting (1997).

Group	Age M (SD)		
	7 years	8 years	11 years
1 (N = 44)	13.66 (2.20)	14.71 (2.20)	17.74 (1.79)
2 (N = 47)	12.32 (1.78)	13.49 (2.40)	15.43 (2.92)
3 (N = 84)	7.44 (2.17)	10.33 (2.72)	14.06 (2.96)
4 (N = 9)	7.44 (2.19)	12.67 (3.12)	15.78 (2.82)

Note. Group 1 = Language normal group; Group 2 = Expressive language impairment group; Group 3 = Expressive/receptive language impairment group; Group 4 = Complex specific language impairment group (poor expressive language skills but lower receptive language skills [below the 10th percentile]). ICD-10 = International Statistical Classification of Diseases and Related Health Problems (10th rev.).

Figure 2. The average trajectories for Receptive Language skills across the three age points for the language-normal group (Group 1), expressive language impairment group (Group 2), expressive/receptive language impairment group, (Group 3), and complex SLI group (Group 4).
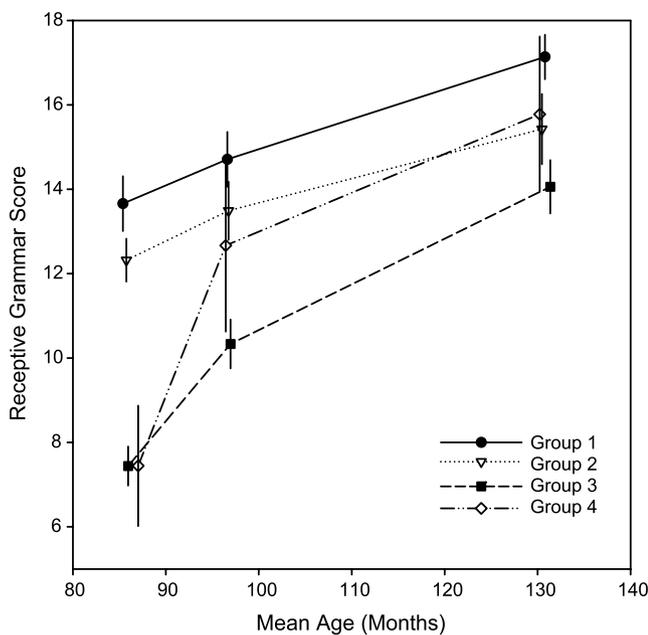
was also a significant difference in overall language status (intercept) for the different groups, $F(3, 180)= 74.09$, $p < .0001$, $T^2 = 0.28$. Finally, the Age × Group interaction was also significant, $F(3, 364) = 23.86$, $p < .0001$, $T^2 = 0.028$. A follow-up test to examine the interaction showed that of the six contrasts, Groups 1 and 2 were not significantly different in the growth rate and Groups 3 and 4 were not significantly different. All other contrasts in the growth rate were significant.

Although the mixed model provided some evidence of reliable differences between some of the subgroups, this fixed-model effect size also showed that these significant between-group differences resulted from very small effect sizes as reflected in the T^2 values of 0.04. This estimate of T^2 used a method described by Winer, Brown, and Michels (1991; pp. 406–407) and assumes a fixed effect of analysis of variance (ANOVA), whereas the significance test was based on a mixed model. To date, there is no method available for computing effect sizes for mixed models; however, the fixed-effect approximation should be quite close. This constitutes a small effect (Cohen, 1988), accounting for only 4% of the variance. Thus, these findings underscore the relative homogeneity of the growth characteristics of these subgroups. The groups did differ with respect to their overall severity as shown by the significant main effect and the large effect size. Taken together, these results indicate that most of

the differences between the subgroups are found at the general level of language proficiency and that, except for a small amount of the variation, the progress across time for these four subgroups is similar.

Receptive Language Trajectories for the Rapin and Allen (1987) Classification System

The mean scores for each of the six Rapin and Allen (1987) clusters are presented in Table 4, and the patterns of language growth for each of the clusters are shown in Figure 3. The pattern is similar to that detected in the earlier analysis in that each of the groups follow a similar incremental increase in language skills across time. A mixed-model ANOVA was again used to test whether the clusters differed with respect to receptive language status across the three observational intervals. This analysis showed that across all groups, there was a significant effect of age, $F(1, 362) = 377.42$, $p < .0001$, $T^2 = 0.19$. The clusters also differed with respect to their overall language status, $F(5, 178) = 24.34$, $p < .0001$, $T^2 = 0.19$. Furthermore, the Age × Group interaction was significant, $F(5, 362) = 3.15$, $p < .01$, $T^2 = 0.006$. A follow-up test to examine the interaction showed that of the 15 contrasts, 3 were significant at an alpha level of $p < .01$. Specifically, Cluster 4 had a growth rate significantly lower than Clusters 1, 3, and 5.
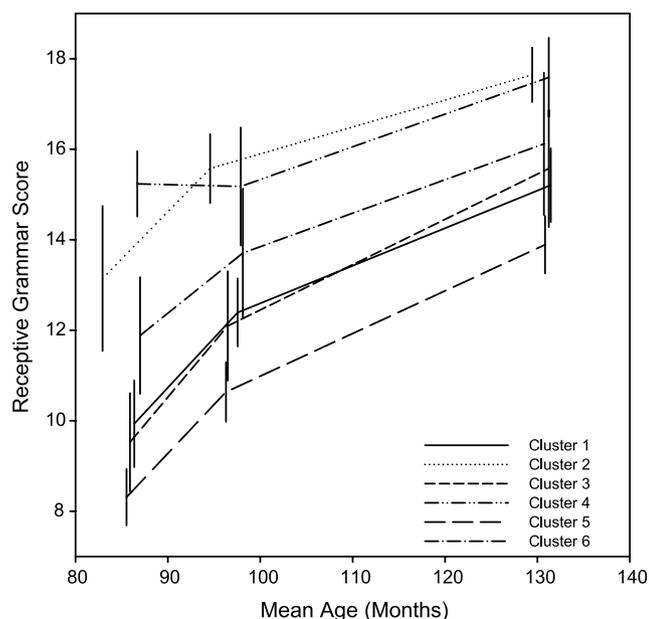
An inspection of the plots in Figure 3 shows that these differences are largely due to the negative growth for Cluster 4 from age 7 to age 8 years. Cluster 4 represents a group with normal language abilities but with

Table 4. Mean (and standard deviation) of the Receptive Language score (TROG) for the Rapin and Allen (1987) classification system at three age points based on clusters identified in Conti-Ramsden et al. (1997).

| | Age | | |
Group	7 years	8 years	11 years
1 (N = 44)	9.93 (3.25)	12.39 (2.54)	15.20 (2.76)
2 (N = 14)	13.14 (3.06)	15.57 (1.45)	17.64 (1.15)
3 (N = 21)	9.52 (2.54)	12.10 (2.83)	15.57 (3.03)
4 (N = 17)	15.24 (1.52)	15.18 (2.74)	17.59 (1.84)
5 (N = 71)	8.31 (2.68)	10.63 (2.83)	14.14 (1.83)
6 (N = 17)	11.88 (2.71)	13.71 (3.00)	16.12 (3.31)

Note. Group 1 = Good articulation but other language skills poor; Group 2 = Poor word reading; Group 3 = General poor language but with good naming vocabulary; Group 4 = Poor articulation and phonology, and normal language; Group 5 = Articulation is fair but performance on all other tests is poor; Group 6 = Good articulation and reasonable expressive language but poor receptive language.

Figure 3. The average trajectories for Receptive Language skills across the three age points for the profile of Conti-Ramsden et al. (1997) classification system containing the following six groups: (1) Good articulation but other language skills poor; (2) Poor word reading; (3) General poor language but with good naming vocabulary; (4) Poor articulation and phonology, and normal language; (5) Articulation is fair but performance on all other tests is poor; and (6) Good articulation and reasonable expressive language but poor receptive language.

was fit to a linear model, 30.3% of the total variance was accounted for by a single set of intercept and slope parameters. Of the remaining variance, nearly 61.6% (100% − 30.3% = 69.7%) was due to individual differences in intercepts, and the remaining 38.4% may be attributed to the individual difference in the growth rate and the variance unaccounted for by the linear model. The intercept reflects differences in severity at initial assessment. Although systems concerned with subtypes can include differences in severity, typically this form of individual difference (i.e., in intercept) is viewed as evidence of dimensional quantitative variation. The individual growth rate differences and the accounted errors can be seen from the samples of growth found in Figure 1. It is within this nearly 26.8% [i.e., 69.7% × (1.000 − 0.616)] of the variance that qualitative differences in the pattern of growth that would contribute to subgroups may be found. It might be possible to assume that if there were no measurement error, an ideal taxonomy with the correct growth function that would capture individual differences in growth patterns would result in an effect size of this magnitude. It is not possible, from these data, to extrapolate this growth function to the population of language learners as a whole, although it is important to point out that whether a child had a clinical level of language difficulty did not account for class membership in the Tomblin et al. (2004) analysis. Thus, the pattern of development in language impairment may indeed reflect that in the typically developing population.

Discussion

This study examined the growth characteristics of language in a large clinical sample of children who have been described in prior research (Conti-Ramsden & Botting, 1999; Conti-Ramsden et al., 2001). Existing work by Conti-Ramsden and colleagues has provided evidence for subgroups based on profiles of language performance across language measures at each time point. Stability of group membership across observation intervals was not found. These findings suggest that this sample of children with language impairment is heterogeneous and this heterogeneity is captured, to some degree, by categories similar to those initially identified by Rapin and Allen (1987). Given the potential heterogeneity of this sample, considerable variability in growth patterns was anticipated across time, and this variability was expected to be associated with the diagnostic subgroups found by Conti-Ramsden and colleagues. Indeed, these authors indicated that 45% of the children in the sample moved across subgroups between Time 1 and Time 2 (Conti-Ramsden & Botting, 1999), and this movement represented substantial changes in the children's

phonological impairment. Thus, their performance on the TROG is close to ceiling. This near-ceiling performance reduces the variance in this group and also the amount of change over time. The variance of the measure plays a role in determining whether or not the growth rate for one cluster is significantly different from that of other clusters. The differences between groups with small variances tend to be significant. This may explain why the growth rate for Cluster 4 is significantly lower than for Clusters 1, 3, and 5 but not for Cluster 2, which has a larger variance for the measure at age 7 years. As was found for the earlier subgroups, this analysis revealed that despite there being a significant difference in the development pattern of some of the clusters, the magnitude of these differences as reflected in the T^2 value of the interaction was very small. In contrast, the main effect between the groups was larger, again indicating that the principal dimension distinguishing the clusters was the severity of their receptive language scores at initial assessment.

The data from this study showed that there was a dominant pattern of development consisting of a gradually declining rate of growth for these children with clinical levels of language impairment. When this pattern

profiles. Such heterogeneity would be particularly likely if these subgroups had different etiologies that would have different effects on the rate and pattern of language growth.

It is within this context of the remaining 26.8% variance not accounted for by the dominant model that the Rapin and Allen (1987) and ICD-10 (WHO, 1992) taxonomic systems can be considered, as this would represent the maximum amount of variance that these systems could account for beyond that having to do with severity. Indeed, both systems were able to capture a significant amount of this variance. The estimated effect sizes for these significant results, however, were quite small (.01 and .04, respectively). The failure to obtain a stronger relationship between these systems of subtyping and growth over time may be due to their poor reliability as shown by Conti-Ramsden et al. (2001). If so, one would have to question whether these subgroups are reflective of substantial unique etiological effects. The failure to find strong evidence for systematic differences in growth for these taxonomic systems could also be explained by the fact that their differences are due to nonlinear aspects of growth that were not incorporated into this analysis. An examination of Figures 3 and 4 shows, however, that the shapes of the group means across time are quite similar both within and between the two taxonomic systems. Therefore, it seems unlikely that the use of a linear model for analysis obscured important group differences.

The substantial similarity in growth trajectories between individuals and within groups formed by existing diagnostic systems is similar to other studies that have recently examined growth in school-aged children with and without specific language impairment. The findings of Rice and colleagues (Rice, 2004; Rice, Wexler, & Hershberger, 1998) and Tomblin et al. (2004) suggest that the growth trajectories of measures of both vocabulary and grammar in children with SLI were similar to those of typically developing children. Indeed, the main differences between groups in both studies was their early language skills, not the pattern of their subsequent performance. As Rice (2004) noted, the acquisition of different aspects of the language system appears to be explained in the first instance by the timing of the onset of acquisition, but subsequently, the growth trajectories follow the same course in all children, although the exact pattern may differ for different speech and language skills. In general, this would support the notion that these children are delayed but not qualitatively different from typically developing children, supporting a finding in the literature that goes as far back as the work of Morehead and Ingram (1976).

The conclusion from the present data is that, in terms of the models of change proposed by Leonard (1998) and discussed in this article, the pattern appears to be similar for the group as a whole. Although some individual children may appear to catch up and others appear to fall behind, the overall pattern appears to be one of consistency. These data do not allow direct comparison with a comparable group of typically developing learners. Those that seem to have normal language scores at later stages in the cohort could hardly be classified as typically developing learners, given the fact that they had been receiving special educational provision at the initial stages of the original Conti-Ramsden Manchester Language Study. However, data from the Iowa Child Language Study (Tomblin & Zhang, 2006) suggest that this pattern of slowing trajectory, arguably to asymptote, is characteristic of all children (Zhang & Tomblin, 2004). It is true that these findings reflect performance on a specific set of standardized language measures, and there may be different patterns for other skills that are not included. It is also true that the patterns of change are different for some individuals but, across the group as a whole and at least as far as receptive language skills are concerned, it appears that the tracking hypothesis better accounts for the data for the receptive language skills of the group as a whole than does either the deterioration or plateau hypotheses. The same pattern might not be found for speech skills or expressive language skills.

Methodological Limitations

The information obtained from this study must be understood within the context of the limitations inherent in the data set used. As noted initially, this study represents a secondary analysis of data that had been collected for a different purpose than that of growth curve analysis. The value of this data set derives from its extensive sampling of a clinically served population. This being said, certain features of this data set were not ideal.

Limited repeated observations. To obtain effective trajectories, it is important to consider the number of observations and the intervals between the time points. Growth curve modeling involves the estimation of parameters in the model. In the case of growth analysis, a minimum of three time points is needed to estimate the intercept and slope. With three observation intervals, it is not possible to test more complex models that contain nonlinear effects or to test for heterogeneity of growth parameters. It is also desirable that the intervals over which the observations are made allow for sufficient time for change to occur. Willett, Singer, and Martin (1998) talk about "sensibly spaced intervals." In the present data sets, there was only 1 year between the performance of the children between 7 and 8 years of age (i.e., the first two data points) in terms of both age and

ability. Data that spanned growth over a larger range of development would have been desirable.

Unidimensional representation of language. In addition to a limitation in the number of observations, the language change data were restricted to receptive language skills. These skills are undoubtedly salient in terms of the long-term outcome for children with language impairment; however, in general, other characteristics may be related to language change that play a significant part and that could be monitored in the same way. For example, it is possible that comparable expressive language trajectories for grammar or vocabulary might be characterized by greater variability, although such an interpretation would be called into question by the data from the Iowa Child Language Study (Tomblin & Zhang, 2006). This analysis was not possible from the available data.

An alternative to this approach would be a person-oriented approach. The *person approach,* as described by Von Eye and Bergman (2003), assumes that multiple and somewhat unique factors influence development and, therefore, the individual is also part of a system and cannot reasonably be seen as being outside the context of that system. To accomplish such person-level analyses, it is necessary to have multiple measures of the traits of interest in order to perform cluster analysis with manifest variables or latent class analysis. The data available for this study did allow for such analyses.

Attrition. It is important to comment on the attrition in the study in the process of identifying sufficient cases to plot the trajectories. In doing this, the analysis effectively lost approximately 25% of the participants. Of course, if this dropout was random, it need not be a problem. However, this explanation is unlikely to be the case, and it may be that this would affect the outcomes. For example, those children with the poorest language skills might have been more likely to refuse specific tasks than children with less severe language impairments. Their removal from the analysis, consequently, would have reduced the number of a specific subset of children, perhaps those with language difficulties who were more closely associated with lower nonverbal IQ scores.

Conclusions

The present data support the contention that that there is unlikely to be a set of diagnostic elements that can provide a mutually exclusive classification system at any one given time (Bishop, 1994). Moreover, our results suggest that whether this is possible or not is probably the wrong question. Rather, it is necessary go beyond the notion of disorders at single time points and see them as

being on temporal continua. The challenge is to identify predictive models that can be replicated with sufficient statistical power to sustain the confidence of researchers and clinicians. The data presented here represent a first step in this process.

This analysis suggests that secondary analysis, which has not been a feature of the literature in the field of speech-language pathology, offers considerable potential for addressing salient research questions especially by using different cohorts to address the same issues. Given the consistency of the results present in this article, it could be anticipated that these findings would hold over other cohorts. However, it would be worth going back to those cohorts that are large enough to provide comparable data, namely those in Dunedin, New Zealand, and Ottawa, Carleton, Canada. It would also be appropriate to test for theoretically and clinically derived trajectories in the data to examine whether there are meaningful latent classes, which can be identified and again replicated across cohorts and with different composite measures. This will lead to further discussions of the dimensionality of language impairment, which can only be addressed in large-scale cohort studies.

References

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

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., & Lancee, W. (1996). Long term consistency in speech/language profiles. 1: Developmental and academic outcomes. *Journal of American Academy of Child Psychiatry, 35,* 804–825.

Bishop, D. V. M. (1982). *The Test of Reception of Grammar* (2nd ed.). Manchester, United Kingdom: University of Manchester.

Bishop, D. V. M. (1994). Is specific language impairment a valid diagnostic category? Genetic and psycholinguistic evidence. *Philosophical Transactions of the Royal Society of London, 346,* 105–111.

Bishop, D. V. M., Bright, P., James, C., Bishop, S. J., & van der Lely, H. K. J. (2000). Grammatical SLI: A distinct subtype of developmental language impairment? *Applied Psycholinguistics, 21,* 159–181.

Bishop, D. V. M., & Rosenblum, D. (1987). Classification of childhood language disorders. In W. Yule & M. Rutter (Eds.),

Language development and disorders (pp. 16–42). Oxford, England: MacKeith.

Botting, N., & Conti-Ramsden, G. (2000). Social and behavioral difficulties in children with language impairment. *Child Language Teaching and Therapy, 16,* 105–120.

Botting, N., Faragher, B., Simkin, Z., Knox, E., & Conti-Ramsden, G. (2001). Predicting pathways of specific language impairment: What differentiates good and poor outcomes. *Journal of Child Psychology and Psychiatry, 42,* 1013–1020.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Conti-Ramsden, G., & Botting, N. (1999). Classification of children with specific language impairment: Longitudinal considerations. *Journal of Speech, Language, and Hearing Research, 42,* 1195–1204.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42,* 741–748.

Conti-Ramsden, G., Botting, N., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: Outcomes at 11 years of age. *International Journal of Language and Communication Disorders, 36,* 207–219.

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language, and Hearing Research, 40,* 765–777.

Dollaghan, C. A. (2004). Taxometric analyses of specific language impairment in 3- and 4-year old children. *Journal of Speech, Language, and Hearing Research, 47,* 464–475.

Elliott, C. D. (1983). *British Ability Scales.* Windsor, United Kingdom: nferNelson.

Goldman, R., & Fristoe, M. (1986). *Goldman–Fristoe Test of Articulation.* Circle Pines, MN: American Guidance Service.

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., et al. (1999). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, and Hearing Research, 42,* 744–761.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2,* 389–398.

Leonard, L. (1998). *Children with specific language impairment.* Cambridge, MA: MIT Press.

Morehead, D. M., & Ingram, D. (1976). The development of base syntax in normal and linguistically deviant children. In D. M. Morehead & A. E. Morehead (Eds.), *Normal and deficient child language* (pp. 209–238). Baltimore: University Park Press.

Rapin, I. (1996). Developmental language disorders: A clinical update. *Journal of Child Psychology and Psychiatry, 37,* 643–656.

Rapin, I., & Allen, D. A. (1987). Developmental dysphasia and autism in preschool children: Characteristics and subtypes. In J. Martin, P. Fletcher, R. Grunwell, & D. Hall (Eds.), *Proceedings of the first international symposium on specific speech and language disorders in children* (pp. 20–35). London: Afasic.

Raven, J. C. (1986). *Coloured Progressive Matrices.* London: H. K. Lewis.

Renfrew, C. (1991). *The Bus Story: A test of continuous speech.* Oxford: Author.

Rice, M. L. (2004). Growth models of developmental language disorders. In M. L. Rice & S. F. Warren (Eds.), *Developmental language disorders: From phenotypes to etiologies* (pp. 207–240). Mahwah, NJ: Erlbaum.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1412–1431.

Richman, N., Stevenson, J., & Graham, P. J. (1982). *Pre-school to school: A behavioral study.* London: Academic Press.

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

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Disorders, 41,* 407–418.

Tomblin, J. B., Records, N., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 4,* 1245–1269.

Tomblin, J. B., Zhang, X., Catts, H., Ellis Weismer, S. E., & Weiss, A. (2004). Dimensions of individual differences in communication skills among primary grade children. In M. Rice (Ed.), *Developmental language disorders: From phenotypes to etiologies* (pp. 53–76). Mahwah, NJ: Erlbaum.

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research, 49,* 1193–1208.

van der Lely, H. (2003). Do heterogeneous SLI deficits need heterogeneous theories? SLI subgroups, G-SLI, and the RDDR hypothesis. In Y. Levy & J. Schaeffer (Eds.), *Towards a definition of specific language impairment* (pp. 109–134). Mahwah, NJ: Erlbaum.

van der Lely, H., Rosen, S., & McClelland, A. (1998). Evidence for a grammar-specific deficit in children. *Current Biology, 8,* 1253–1258.

Von Eye, A., & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-orientated approach. *Development and Psychopathology, 15,* 553–580.

Willett, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology, 10,* 395–426.

Wilson, B. C., & Risucci, D. A. (1986). A model for the clinical quantitative classification—Generation 1—Application to language disordered preschool children. *Brain and Language, 27,* 281–309.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design*. New York: McGraw-Hill.

World Health Organization. (1992). *International Statistical Classification of Diseases and Related Health Problems* (10th rev.). Geneva: Author.

Zhang, X., & Tomblin, J. B. (2004, June). *Lack of individual difference in the language growth rate from kindergarten to eighth grade*. Presentation at the Society for Research in Child Language Disorders, Madison, Wisconsin.