The natural history of speech impairment of 8-year-old children in the Avon Longitudinal Study of Parents and Children: Error rates at 2 and 5 years

SUE ROULSTONE[1,2], LAURA L. MILLER[3], YVONNE WREN[1,2], and TIM J. PETERS[3]

1. University of the West of England, Bristol, UK
2. Speech and Language Therapy Research Unit, Bristol, UK
3. University of Bristol, UK

Correspondence: Sue Roulstone, Professor of Speech & Language Therapy, Speech & Language Therapy Research Unit, Frenchay Hospital, Bristol, BS16 1LE, UK. Tel: +44 117 9186529. Fax: +44 117 970 1119. Email: sue@speech-therapy.org.uk

Running head: Natural history of speech impairment

Keywords: phonology, ALSPAC, natural history, speech impairment.

Abstract

This prospective longitudinal population study observed the speech of 741 children at the ages of 2, 5 and 8 years. At the age of 8, 132 children were categorised as speech impaired. There was strong evidence of differences between the case and control groups in speech sound error rate at the two earlier age points. The pattern of the proportion of errors was similar for cases and controls. There was evidence of a strong relationship between the child's error rate and expressive language at age 2 and between the child's error rate and sentence length at age 5. In multivariable analyses, adjusting for expressive language, parent's social status, maternal age, gender and child's exact age, the increase in odds of being a case as the proportion of errors increased disappeared at 2 years but remained at 5 years. At 5 years, the proportion of speech errors was predictive of ongoing speech errors at the age of 8 years; the adjusted odds of having speech errors at 8 years increased by between 21 and 44 % with every 10% rise in the proportion of errors in the target patterns.

Cohort studies are the design of choice to investigate a range of epidemiological questions about health and disease and health related conditions. Typically they involve tracking a population over a period of time and are thus prospective and longitudinal (Bhopal, 2002). Rather than starting with participants who already have evidence of the disease or condition, the starting point for population based cohorts is a healthy population. A population based cohort study provides a context for observing the emergence of the condition in the context of a range of exposures, that is events that appear to be associated with particular outcomes, namely the disease or condition (Bhopal, 2002). Population cohorts are particularly useful for observing the natural history of conditions such as speech and language impairments in children, where the cohort is established prospectively, before the children have been identified with any speech and language impairment.

The natural history of any condition describes the process by which the condition occurs and progresses. In the context of disease, one would expect to see descriptions of the changes that lead from a healthy state to a disease state, from “prepathogenesis” where the individual may be vulnerable to the disease but have no unhealthy symptoms through early and late phases of a disease (Valanis, 1999, p. 270). The possibility of identifying a presymptomatic phase in developmental conditions is debateable (Law, Boyle, Harris, Harkness, & Nye, 1998). Furthermore, as Gruber et al. (2003) note, the possibility of isolating a natural history path that is truly without interventions is virtually impossible. Nonetheless the aim of identifying particular phases and stages in the progression of speech and language impairments remains important to our understanding of the conditions, how and when to intervene and with what kind of program. An understanding of natural history also enables us to evaluate the added value of our interventions (Gruber et al., 2003). In addition to the difficulties of identifying a pre-symptomatic phase, there are other difficulties. Law et al. (1998) for example note the changing presentation of speech and language difficulties over time and the lack of a linear relationship between the original presentation of the difficulty and its later outcome; these make the investigation of the natural history a complex challenge. However, Weiss (1996) remarks that this is one of the advantages of the longitudinal study since it can track the continuities and discontinuities of a condition and identify events and exposures that are associated with those changes.

The concept of natural history of speech and language impairments is one that has been present in our literature over the years but rarely explored explicitly. In their systematic review of the natural history of speech and language delay, Law et al. (1998) noted that whilst there were many follow-up studies that had investigated the prognosis and longer term outcomes of speech and language delays, relatively few had investigated natural history. The review identified only twelve studies up to the point of publication in 1998 where there had been no specific treatment. Only three of the studies focused entirely on speech (Bralley & Stoudt, 1977; Fiedler, Lenneberg, Rolfe, & Drorbaugh, 1971; Renfrew & Geary, 1973) and one included both speech and language measures (Felsenfeld, Broen & McGue,1992); the other eight focused on language. The report by Fiedler et al. does not differentiate between the speech and language components of the children’s delay or provide an analysis of the kinds of errors made by the children who were identified as delayed. Bralley and Stoudt (1977) assessed 120 children entering their first grade and identified 60 who "misarticulated" at least one phoneme. They then assessed these 60 children at the start of each subsequent year noting a gradually decreasing error rate; by Grade 5 only 13 of the participants showed articulation errors and two participants contributed more than half of the errors made. The highest number of errors were made in /l/ constant clusters (16 errors), /z/ (9 errors), /θ/ and /s/ consonant clusters (each showing eight errors). None of these children had received intervention. Bralley and Stoudt conclude that although the majority of children appear to have resolved articulation errors by fifth grade, intervention was still indicated for a significant few. The study by Felsenfeld et al. (1992) suggests that speech errors can still be apparent in adulthood. They followed a cohort of children (n = 24) who had presented with moderately severe phonological impairments. As adults they were found to make significantly more errors than a control group who had also been followed over the same time period. Whilst the speech of the control group was error free, those who had impairments as children showed residual segmental errors, notably, with /r/ and sibilants /s, z/. Approximately 20% of the impaired group also showed occasional simplification of consonant clusters; also noted were differences in the suprasegmental features such as stress and intonation although these were not compared quantitatively with the control group.

More recently, a study by McIntosh and Dodd (2008) assessed 62 children aged between 25 and 35 months. Ten of these children were assessed three times within a twelve month period on the Toddler Phonology Test (TPT) for the first two and then on the Diagnostic Evaluation of Articulation and Phonology (DEAP, Dodd et al., 2002). At baseline, five of these ten children performed more than one standard deviation below the mean of the sample regarding the percentage of consonants correct (PCC). When all ten children were assessed 6 to 11 months later on the on the DEAP, they found significant correlations between the three assessment points. However, at the final follow-up, all the children performed within the normal limits as defined by DEAP. The study also analysed data on the type of errors made and found that two of the children who showed atypical errors at baseline continued to show atypical errors at age 3. McIntosh and Dodd concluded that the children's PCC scores did not predict their outcome on the gold standard assessment and that qualitative measures were more sensitive. The authors do indicate that larger numbers are needed to determine the predictive validity of qualitative and quantitative measures. Given the small number of children in the study and the absence of any who did not fail the DEAP at follow up, it would indeed be presumptive to conclude that quantitative measures at the age of 2 years cannot predict outcome at the later time point.

In addition to the small number of longitudinal studies of children with speech impairments, there are numerous studies of speech acquisition, ranging from the early emergence of speech sounds through to a child's acquisition of clusters (see McLeod, 2009 for a review of studies of children's speech acquisition). Many report on the individual differences between children in the age at which they acquire sounds and, although the bases of these differences are far from confirmed (Stoel-Gammon, 2007), a number of candidate variables have been identified. The early study by Renfrew and Geary (1973) followed children over a six month period, with the first assessment taking place between 2 and 5 months after school entry. Factors that predicted which children had improved included their ability to imitate sounds spontaneously, to discriminate between like-sounding words and to repeat simple tongue-twisting phrases. Williams and Elbert (2003) in a longitudinal study of 5 late talkers identified atypical error patterns (such as d/h and dʒ /t) and greater sound variability as potential markers of long-term phonological delay. Considering broader demographic factors, there have been reports regarding the range of factors associated with speech impairment. Results are sometimes apparently conflicting. For example, the association between speech impairment and male gender is a common finding (Campbell et al., 2003; Keating, Turrell & Ozanne, 2001). Yet Dodd et al. (2003) found this association only in their older age groups. Associations with levels of maternal education were found by Campbell et al. but not by Roulstone et al. (2002). As Campbell et al. (2003) note, their study was the first to provide odds ratios regarding the risk of speech impairment relative to the putative variables.

One of the key variables likely to be associated with the individual differences that are observed, particularly in the cross sectional studies where children are assessed at the same chronological age, is the complexity of the children's expressive output at the time of the phonological assessment. Roulstone et al. (2002) found that the proportion of speech errors observed in a single word sample from 2-year-old children was related to their reported sentence length showing a decrease in errors as sentence length increased. Shriberg et al. (2000) found a significant mediating effect of language on speech in the context of children with a history of hearing loss. Variations noted in the rate of error or the range of sounds used may be associated in part to differences in the child's expressive language ability since there is an identified relationship between the children's developmental expressive language stage and the proportion of errors observed (Roulstone et al., 2002). More recently Vogel Sosa and Stoel-Gammon (2006) have observed changes in the levels of variability in a longitudinal study of four children. They noted that although children's early productions were highly variable, this was followed by a period of relative stability and subsequently a further period of instability, with intra-word variability peaking at the point of onset of word combinations. There have been a number of studies of the phonologies of late talkers or of children with specific language impairment (Paul & Jennings, 1992; Rescorla & Ratner, 1996; Williams, & Elbert, 2003) who find that the late talking children have less well developed phonologies than either typically developing peers or than children who later catch up with their peers. However, Schwartz, Leonard, Folger and Wilcox (1980) found very few differences in the phonology of the language disordered children and that of the younger children matched for mean length of utterance, sex and sensori-motor intelligence. In a follow-up study to Rescorla and Ratner (1996), Roberts et al. (1998) showed that the late talkers had caught up with their peers in vocalisation rate but were still behind in their phonetic inventories, percentage of consonants correct and intelligibility. Similarly Pharr et al. (2000) examined syllable structure in children with expressive specific language impairment (SLI-E) at 24 and 36 months. They found that typically developing children produced more consonant clusters at 24 months than the children with SLI-E at 36 months and that children with SLI-E vocalised less than typically developing peers at both ages. Again, adjustments were not made for differences in expressive language stage or output. The results from these studies suggest that it is important to take the language level of the child into account when assessing their phonology.

A consideration of the literature regarding speech development highlights the varying sample size and type of data collected on the children's speech. In fact, as McLeod (2009) notes, there is usually a trade-off between the level of detail of data and the number of participants. McLeod (2009) identified eleven reports of speech acquisition, including two which also provide cross-sectional data. The noted trade-off is interestingly reflected in the longest and shortest study of those identified by McLeod: Oller et al (1999) studied 3400 children in a mixed longitudinal/cross sectional study over a period of 2 months – the shortest period and the largest sample; on the other hand Selby Robb and Gilbert, (2000) study four children over a period of 21 months – the longest longitudinal study identified, with the smallest sample. This is likely to be due to feasibility issues. With larger samples, the research time (and funding) available for each individual is often relatively small thus limiting the level of detail that can be achieved regarding the individual child's speech. The smaller studies on the other hand, can provide in-depth investigations which are crucial to the development of theoretical propositions. The studies with the larger sample sizes are then needed to test the hypotheses and investigate the broader applicability of the theory.

In conclusion, studies of children's speech acquisition have generated considerable knowledge about the order of emergence of speech sounds, error patterns and phonological processes. However, there are few studies which track children with speech impairment longitudinally. Typically these are either small scale, with clinical populations or contain limited information about the child's emerging phonology. Evidence regarding factors associated with persisting speech impairments has provided contradictory results and rarely is adjustment made for a child's emerging expressive language.

The study reported in this paper uses data from the Avon Longitudinal Study of Parents and Children (ALSPAC), a large prospective population-based study which has collected speech data on children at three time points between the ages of 2 and 8 years. The paper will first provide some general background information about ALSPAC, the population study from which the data in this paper are derived, then will describe the particular methods within ALSPAC of relevance to this paper. The aim of the paper is to examine the error patterns of children with persistent speech impairment at 8 years when they were aged 2 and 5 in order to better understand the natural history of speech impairment. Research questions included:

Are the error patterns of children with persisting speech impairment at 8 years distinct from control children at earlier time points, specifically at the ages of 2 and 5 years?

Does the pattern of speech errors change if the child's expressive language output is taken into account?

Do the child's error patterns at 2 and 5 years predict outcome at 8 years?

The Avon Longitudinal Study of Parents and Children (ALSPAC)

Before focusing on the speech data, it is useful to provide some general information about ALSPAC. During 1991 and 1992, 14,541 mothers enrolled in ALSPAC as they registered their pregnancy within the geographical area then known as Avon in the southwest of the UK. From these women's pregnancies, 13,988 children were alive at one year after birth, which included multiple births. The overall aim of the study is to investigate gene-environment relationships to identify optimal pathways to wellbeing for individual or given genotypes (Golding, Pembrey, Jones & the ALSPAC study Team, 2001). The main data collection technique for the whole study has been via postal survey: the mothers completed four questionnaires before their babies were born and approximately annually thereafter, with 16 completed by the time the child was aged 13 years. The children also completed their own questionnaires. Since the children were aged 7 years they have also been invited to attend for direct assessment on seven subsequent occasions (the 'Focus' clinics). In addition, a 10% subsample, chosen at random from the last six months of the cohort and known as the 'Children in Focus' were seen for direct assessment every 6 months from the age of 4 months until the age of 6 years.

Data collected has included biological samples (for example, blood, teeth, toenails, placenta), information about the child, mother and partner, about the general environment of the family home and psychosocial aspects of the family. Information about the children's cognition, speech and language, social and educational development has been collected by parent report and by direct assessment. Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees.

Method

The data presented in this paper is part of a larger analysis of speech data in ALSPAC. Two associated studies are currently examining the origins, trajectories and impact of stuttering and of speech impairment. In the latter, the term speech impairment is used as a generic overarching term for those impairments which impact on the intelligibility of a child's spoken output, irrespective of the origin, although within the study, analyses have focused on subgroups defined by the types and combinations of errors children make in terms of distortions, substitutions and omissions. Children who stutter are not excluded from the sample of children with speech impairment.

Participants

In this paper we focus on participants who attended the Children in Focus clinics at 2 and 5 years and who subsequently also attended the Focus clinic at 8 years. Figure 1 shows that, of the 7488 children who attended the Focus at 8 years clinic, 7390 completed the speech assessment component; of those 741 attended and completed the speech assessments at the Children in Focus clinics and 2 and 5 years. Of the included children 391 (52.8%) were boys and 350 (47.2%) were girls; nearly 60% of the children had a parent working in non-manual occupations (also known white collar) (OPCS, 1991).

The children attended the clinic sessions with their parents for up to 3 hours at each visit. These clinics were held in university premises. During the visit each child completed a circuit of assessments including height and weight, vision, hearing, dietary records, dental examination as well as psychological and speech and language testing. The order of the assessments that children carried out was adjusted to take account of possible order effects. Full details of all the assessments conducted at each clinic are available on the study website: (http://www.bristol.ac.uk/alspac)

Assessment at 2 years

At 25 months, approximately 20 minutes was allocated for speech and language assessments. At the clinic, parents completed a short questionnaire about their child's expressive output indicating their child's current level of expression. Box 1 shows the question and options available to the parent. In addition children completed a single word object naming task adapted from Pagel Paden et al. (1987). At the time of planning the assessments in 1991, this was one of the few word lists available that had been developed with consideration for emerging phonologic classes and syllabic structures of pre-school children and which also had data regarding performance levels for 2-year-olds. Real objects were hidden in a bag and the children were encouraged to take out the object and label it. If children did not spontaneously name the object, the tester labelled the object and encouraged the child to imitate. Following Pagel Paden et al., we agreed that for 2-year-old children, it was important to provide a short task rather than attempt a full phonological assessment, particularly in view of within-word variability in children of this age. Testing the number of errors rather than attempting to identify any particular pattern of substitutions or simplifications avoids this particular pitfall (Pagel Paden et al., 1987). Adaptations were made to the original Pagel Paden et al., word list to ensure culturally relevant vocabulary and to take account of the UK accent in which the postvocalic /r/ is not mandatory. The words selected were designed to sample a minimum of ten errors in the following phonologic classes: velars, postvocalic singleton consonants, liquids, fricatives and consonants clusters, subsequently referred to as the target patterns. The final list of 16 words showing the potential number of errors sampled is in Appendix A. Testers noted the number of opportunities within each target pattern (dependent upon the words actually named by the child) and the children's errors (error pattern). Omissions, substitutions or distortions would all be counted as an error. In each analysis, only children who produced enough words to provide ten samples of a target pattern were included.

Insert Box 1 about here

Assessment at 5 years

A similar process and period of time was allocated for children at the 5 year assessment. The children's expressive language was assessed using the Renfrew Bus Story (Renfrew, 1995), a test of narrative from which is derived the child's average longest sentence length. A further adaptation was made to the single word naming assessment to ensure that the words and target patterns sampled were sufficiently difficult for this age group so that a ceiling effect was unlikely. To achieve this, we included a high proportion of consonant clusters and multisyllabic words. The final word list and target patterns sampled are provided in Appendix B. The same process was used, as above, for noting opportunities and errors; again, only those where there were at least ten opportunities for the target pattern were included.

Assessment at 8 years

When the children attended at the age of 8 years, there was no specific assessment of their speech output. Instead speech samples were collected during an expressive language task based on the Weschler Objective Language Dimensions (Rust, 1996). The expressive component consisted of three picture description tasks. The first required children to describe a picture of a complex town scene; the second required children to provide instructions based on the town scene of how to get from one part of town to another; in the third children were asked to describe what was happening in a sequence of pictures of a torch battery being changed. During this activity, assessors (who were mostly speech-language pathologists (SLPs) and occasionally psychologists) were asked to make observations about the child's speech, fluency, and voice. From these observations, a child was classified as speech impaired if their speech contained errors (substitutions, omissions or distortions). Single errors and dialectal variations were not included. A detailed analysis based on error types is the subject of a separate study (Wren et al., in preparation).

Statistical methods

Demographic data were used to summarise the case and control groups. These two groups were then summarised according to the proportion of errors for each target pattern (% of incorrect postvocalic, velar, fricative and liquid consonants and consonant clusters). These speech error variables at 2 and 5 years were then adjusted (by dividing the score by 10) such that a single unit change relates to a change of 10 percentage points for ease of interpretation in the regression models. Univariable linear regression was used to assess evidence of a difference without adjusting for other factors to allow interpretation of the difference in proportion of errors between cases and controls. For each target pattern at 2 years the relationship with reported expressive language was assessed using linear regression. Similarly at 5 years the relationship with longest average sentence was assessed. Multivariable logistic regression was undertaken adjusting for potential confounding factors (parental social class, maternal age, child's exact age at the relevant clinic and the 8 year clinic, expressive language (2 years only) and average longest sentence (5 years only)). This allowed estimation of the change in odds of being in the case group compared to being in the control group or control for a 10-percentage points increase in the proportion of errors.

Results

Participants and missing data

A total of 741 children presented data at all three time points and thus were included in the study. Analysis of those with missing data suggests that these were children with poorer expressive language as measured by talking stage (single words, 2 words together, 3/4 word sentences: p=0.036), whose parents had manual jobs (p<0.001), and who had a younger mother (p<0.001). Within each clinic there were children who for various reasons failed to complete the speech data assessments, including failure or inability to cooperate or shyness. Of those children who had data at all three time points, 132 were identified at 8 years as making speech errors during the picture description task. In the analysis, the remaining 609 act as the comparator group.

Insert figure 1 about here

Within the speech assessments at 2 and 5 years, children sometimes failed to produce a word or several words. As indicated above, children's data were only included in the analysis of each target pattern where they had produced at least 10 attempts at the target pattern. When those children who completed fewer than ten attempts are excluded, the proportions of cases versus controls is about the same, indicating that by excluding these children we are not losing predominantly case children. A further three children were excluded from the 2 year analysis who were reported by their parents to be only babbling at the time of the 2 year assessment. Numbers for each target pattern are therefore variable. Additionally, there is some missing data for some of the variables used in the multivariable analyses. Table 1 shows the numbers available for the univariable and multivariable analyses.

Insert table 1 about here

The aim of the clinics at 2, 5 and eight years was to see the children within approximately one month window of that age. As Table 2 shows, this is largely achieved and although there is some variation, there is no difference between the case and control children in terms of their age at assessment. The ages of the child's mother were also similar in both groups. The proportion of boys is larger in the case children with nearly 60% of the cases being male and just over 51% of the controls. A slightly higher proportion of the case children had parents working in non-manual occupations. There is no detailed data on attendance for speech-language pathology interventions; however, 28 % of case children and 12% of the control children had attended for assessment.

Insert Table 2 about here

Comparing sound errors in case and control children

Table 3 shows the percentage of errors for each target pattern at 2 years for both cases and controls. This shows a consistent pattern of error patterns in the two groups. That is, the children make fewest errors in postvocalic singleton consonants in both groups, with a progressive increase in velar consonants, then fricative consonants, then liquid consonants, with the highest proportion of errors in consonant clusters. The pattern is slightly changed at 5 years with fricative consonants showing the fewest errors, followed by velar and postvocalic consonants, consonant clusters and finally liquid consonants showing the highest rate of error. Although different from the 2 year pattern, the case and control groups at 5 years look similar. Looking across time, it is clear that the proportion of errors reduces considerably by 5 years with the control group making very few errors particularly for the fricative consonants. Despite the strong patterns that emerge, the large standard deviations at 2 years suggest that there is considerable variation for both cases and controls in the proportion of errors observed; the standard deviations for cases and controls had reduced considerably by 5 years.

Insert Table 3 about here

Univariable linear regression was used to assess evidence of difference between the case and control groups on each target pattern without adjusting for other factors. Table 4 shows strong evidence of a difference between cases and controls at 5 years. Although the differences in the error rate were slightly larger for the two groups at 2 years, these did not provide such strong evidence of a difference; in particular, no evidence of a difference was shown for velar or liquid target patterns.

Insert Table 4 about here

Relationship with expressive language

Before testing to see whether differences between the two groups continued to be apparent when the analyses were adjusted for confounding variables, we firstly wanted to confirm whether or not a relationship existed between the expressive language of the child and their speech status at the time of the phonology assessment and the target patterns of interest. Table 5 shows the mean (and standard deviation) error rate for the various target patterns relative to the child's expressive stage at the age of 2 years. This shows strong evidence of a decreasing error rate in each target pattern as the child's expressive language increases. The pattern of error rate is almost the same as shown in Table 3 with the exception of velar consonants in the context of children using single words: children using single words made fewest errors in velar consonants, whereas those using longer utterances made fewer errors in postvocalic consonants.

At 5 years, the children's expressive language was assessed using the Renfrew Bus Story; thus, the score is a continuous one rather than categorical. Table 6 shows that there is strong evidence of a relationship between the child's longest average sentence and their error rate on all the target patterns except liquid consonants, with a reduction in the proportion of errors as sentence length increases.

Insert Table 5 and 6

Multivariable logistic regression

Previous research has indicated other important associations with the development of children's sound systems, so a multivariable analysis adjusted for parents' social status, maternal age, gender and child's exact age at the relevant clinic. Also, because of the association found above, we also adjusted for reported expressive language (at 2 years) and average longest sentence (from the Renfrew Bus Story at 5 years). In order to examine the contribution of the child's expressive language, a series of multivariable analyses were conducted with and without expressive language as a covariate. Table 7 shows that when adjusting for the above variables (including age specific expressive language of the child), differences in rate of errors between case and control disappeared at 2 years but remained at 5 years; that is, the case children were still presenting with more speech errors at 5 years despite adjustment for potential confounding factors. We conducted the multivariable analyses with and without expressive language in order to examine the contribution of language. This suggests that although expressive language attenuates the relationship between speech sound errors at 2 years and case status at eight years, the language variable is not a major contributor overall (smallest p-value: 0.32). In fact the only covariate which has a noteworthy attenuating effect at 2 years is the mother's age; this was apparent for velar (p=0.043), fricative (p=0.099) and liquid target patterns (p=0.055).

At 5 years, the proportion of errors in the various target patterns is predictive of case status at eight years for all five target patterns. So for example, at 5 years, having adjusted for other variables, there is a 44% increase in the odds of being a case at 8 years with every 10% rise in the proportion of errors made in postvocalic consonants.

Insert Table 7 about here

Discussion

Using prospectively collected data from a longitudinal population study, this paper has examined features of the natural speech history of children with and without persisting speech impairments at the age of 8 years. Although some of the children may have received speech and language therapy for their speech and possibly developmental language delays, this was not co-ordinated or documented within this study. One might expect that intervention would reduce the number of cases identified at the age of 8, although the broad definition of speech impairment used in this study is likely to have counteracted that. The presence of speech data at three time points over an extended period has enabled the tracking of the children's progress. ALSPAC is a large population study which has collected a vast array of data about many features of the children and their families. This paper focused on a small number of social and demographic features in order to understand how they impact upon observed differences in the speech of the children with and without persisting speech impairment. The results are discussed with respect to the consistency of error patterns and the relationship between error pattern and confounding factors in terms of understanding the natural history of speech impairment in children.

Identification of children with speech impairment from the speech samples at 8 years of age produced a high number of cases (18%). Prevalence rates for speech impairment vary widely in the literature from 1.5% (Stewart et al, 1986) to 24.6% (Tuomi and Ivanoff, 1977) and 23% (Jessup et al, 2008). Such variability in part relates to the age being studied but also reflects varying definitions of what constitutes speech impairment. In this study, a broad definition has been used: a child is defined as having a speech problem if they demonstrate perceptible errors in a connected speech sample. This is consistent with the definition used by Shriberg et al. (1997) of speech delayed and questionable residual errors (QRE) where speech delay refers to children who have persisting deletion and substitution errors and QRE is the description used for children who use common clinical distortions such as labialized /r/ and dentalized sibilant fricatives. In the lifespan data of 25 8- year-old children presented by Shriberg et al., they classified eight as either speech delayed or QRE. Using this broad definition, 32% of children were identified as having a discernible speech error in a connected speech sample. This broad definition of speech problems was used in our study because the children were aged 8 when their case status was determined and by this age, the majority of children have developed the full range of speech sounds and sound combinations. The aim was to gain a view of natural history across a broad range of children with ongoing speech impairments at the age of 8 and so a broad definition was deemed appropriate.

The errors observed in this study show a progressive decrease in the proportion of errors relative to the stage of expressive language of the child. The pattern of errors was slightly different across the two age groups, although in both cases, consonant clusters and liquid sounds appear to cause the most difficulty for the children. The patterns found in this study are consistent with other reports. For example Preisser et al. (1988) report the error patterns of 20 children aged 2;2 – 2;5 years; this shows the greatest error rates with liquid consonants and consonant clusters, with reducing error rates observed in fricatives, velars and postvocalic obstruents respectively. Dodd et al. (2003) report that gliding was one of the only processes being used by 10% of children over the age of 5 years in their sample of British English-speaking children, again, consistent with the ALSPAC sample where the highest error rate was observed in liquids, with a mean of 9.6% of the total sample (ie cases and controls) making errors.

In terms of comparing the case and control children, these error patterns follow a similar progression for both groups. Although caution is needed in extrapolating from error rates to acquisition patterns and from group data to patterns for an individual, these data suggest that the apparent order of acquisition is the same for children with and without speech impairments. Until now we have had no indication of the possible progression of sound acquisition in children with speech impairment. This is an important finding for clinicians in terms of deciding the order in which to target patterns for intervention. However, clinicians should be aware of the word list used to elicit these data: the number of words used was small and the words used at the 2 year clinic contained considerable number of consonant clusters. Thus children's facility (or lack of it) with some of the target patterns reflects the context in which they are used. Nonetheless, the data show the wide range present in this population sample. For example some of the 2-year-old children are using three-four word sentences and making only one or two errors with liquid sounds. Other children are using only single words and used no liquids at all. It is also important to note that the analysis produced here shows only the proportion of errors; it does not provide an analysis of what the children produced in lieu of the error. So although both case and control children make fewer errors in velar consonants than they do in fricatives, there may be marked differences in the types of errors used by the two groups. As suggested by McIntosh and Dodd (2008), the error type may be more predictive than the frequency of errors. However, this data provides additional information about the quantitative profiles of children with speech impairment, providing normative data for error rates for children in the different target patterns and not simply the overall PCC. Alongside this we have to consider that the case group was known to be heterogeneous. It may therefore be that despite the apparent consistency of error patterns, there may be a number of subgroups within the case sample. Further analysis of the case group is currently underway. It is also important to acknowledge the considerable individual variation that is apparent for both cases and controls. As one would expect this is greater at 2 than at 5 years, but still apparent even when the child's expressive language stage is taken into account. Furthermore, it is important to remember that the speech samples at 2 and 5 years are small, with between 10-24 opportunities for the various target patterns to occur. Where the number of opportunities is as low as 10, the proportion of errors can only be 0, 10, 20 and so on through 100%, thus limiting the range of possible values.

The multivariable analyses suggest that when adjusted for other variables (that is, expressive language, gender, maternal age, child's age and parent's social status) the differences in the proportion of errors between case and control groups within each target pattern in children aged 2 years is not predictive of persisting speech impairments at eight years. The differences apparent in the univariable analyses are attenuated by the confounding variables and although we found a relationship between the child's expressive language and their error rates across the target patterns in the univariable analyses, only the mother's age seems to be a potent confounder. This could be a chance association. However, it did appear in the analysis of most target patterns. It is therefore likely that mother's age is acting as a proxy for some other factor that is related to both maternal age and case status that we have not accounted for. Given that we know that our case and control groups did not lose proportionately more or less younger mothers, it is difficult to identify what this association might be. As noted above, it is important to note that this is only for the proportion of errors, not for error type. In the older children (that is, at 5 years), the proportion of speech sound errors made by the case children continues to be predictive of ongoing speech impairments at the age of 8 years having adjusted for the child's language and other demographic factors. This suggests that in the younger children the developmental variation and 'noise' explains differences between case and control children more than the error rate in the speech. In the older children however, this early developmental noise has dissipated making the error rate in any of the target patterns a powerful predictor of outcome at eight years.

In conclusion, these data allow us to understand the early natural speech history of children who show persisting speech impairments at the age of 8 years. Whilst the case children do make noteably more errors than the controls at the age of 2 years, the adjusted analyses suggest that the differences between cases and controls in their error rates are explained by factors others than the differences in error rates, so that the errors in the particular target patterns in themselves are not predictive of later case status. By the age of 5 years however, the proportion of errors is a predictive indicator.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. We are particularly grateful to the speech team who collected the speech and language data at the three time points. The UK Medical Research Council, the Wellcome Trust and the University of Bristol provide core support for ALSPAC. This research was specifically funded by the Medical Research Council, UK.

We would also like to acknowledge the contribution of Sue Loader, speech and language therapist, who had the foresight to initiate the collection of speech and language data from the children when they were aged 25 months.

References

Bhopal, R. (2002). *Concepts of epidemiology: An integrated introduction to the ideas, theories, principles and methods of epidemiology*. Oxford: Oxford University Press.

Bralley, R. C. & Stoudt, R. J. (1977). A five year longitudinal study of development of articulation in elementary school children. *Language, Speech, and Hearing services in Schools, 8*(3), 176-80.

Campbell, T. F., Dollaghan, C. A., Rockette, H. E., Paradise, J. L., Feldman, H. M., Shriberg, L. D., Sabo, D. L. & Kurs-Lasky, M. l. (2003). Risk factors for speech delay of unknown origin in 3-year old children. *Child Development, 74*(2), 346-357.

Dodd, B., Holm, A., Hua, Z. & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics and Phonetics, 17*(8), 617-643.

Dodd, B., Zhu, H., Crosbie, S., Holm, A. & Ozanne, A. (2002). *Diagnostic Evaluation of Articulation and Phonology*. London: Harcourt.

Felsenfeld, S., Broen, P. A. & McGue, M. (1992). A 28-year follow-up of adults with a history of moderate phonological disorder: Linguistic and personality results. *Journal of Speech and Hearing Research, 35*, 1114-1125.

Fiedler, M. F., Lenneberg, E. H., Rolfe, U. T., Drorbaugh, J. E., (1971). A speech screening procedure with three year old children. *Pediatrics, 48*(2), 268-76.

Golding, J., Pembrey, M., Jones, R. & the ALSPAC Study Team (2001). ALSPAC – The Avon Longitudinal Study of Parents and Children 1. Study methodology. *Paediatric and Perinatal Epidemiology, 15,* 74-87.

Gruber, F. A., Lowery, S. D., Seung, H-K, & Deal, R. (2003). Approaches to speech/language intervention and the true believer. *Journal of Medical Speech-Language Pathology*, 11 (2), 95-104.

Jessup, B., Ward, E., Cahill, L. & Keating, D. (2008). Prevalence of speech and/or language impairment in preparatory students in northern Tasmania. *International Journal of Speech-Language Pathology, 10*, 364-377.

Keating, D., Turrell, G. & Ozanne, A. (2001). Childhood speech disorders: Reported prevalence, comorbidity and socioeconomic profile. *Journal of Paediatric Child Health, 37*, 431-436.

Law, J., Boyle, J., Harris, F., Harkness, A. & Nye, C. (1998). Screening for speech and language delay: A systematic review of the literature. *Health Technology Assessment, 2*(9).

McIntosh, B. & Dodd, B. J. (2008) Two-year-olds phonological acquisition: Normative data. *International Journal of Speech-Language Pathology*, 10(6), 460-469.

McLeod, S. (2009). Speech sound acquisition. In J. E. Bernthal, N. W. Bankson & P. Flipsen Jnr (Eds.), Articulation and phonological disorders: *Speech sound disorders in children (6th ed.)* (pp. 63-120 + 385-405). Boston, MA: Allyn & Bacon.

Preisser, D. A., Hodson, B. W. & Paden, E. P. (1988). Developmental phonology 18-29 months. *Journal of Speech and Hearing Disorders, 53*, 125-130.

Office of Population Census & Surveys (1991). Standard Occupational Classification. London: Her Majesty's Stationery Office, 1991.

Oller, D. K., Eilers, R. E. Neal, A. R. & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders, 32,* 223-245.

Pagel Paden, E., Novak, M. A., & Beiter, A. L. (1987). Predictors of phonologic inadequacy in young children prone to otitis media, *Journal of Speech and Hearing Disorders, 52*, 232–242.

Paul, R. & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech and Hearing Research, 35*, 99-107.

Pharr, A. B., Ratner, N. B. & Rescorla, L. (2000). Syllable structure development of toddlers with expressive specific language impairment. *Applied Psycholinguistics, 21*(4), 429-449.

Renfrew, C. E. & Geary, L. (1973). Prediction of persisting speech deficit. *British Journal of Disorders of Communication, 8*(1), 37-41.

Renfrew, C. E. (1995). *Bus Story: A test of narrative speech*. UK: Winslow Press.

Rescorla, L. & Ratner, N. B. (1996). Phonetic profiles of toddlers with specific expressive language impairment. *Journal of Speech and Hearing Research, 39*, 153-165.

Roberts, J., Rescorla, L., Giroux, J. & Stevens, L. (1998). Phonological skills of children with specific expressive language impairment (SLI-E): Outcome at age 3. *Journal of Speech, Language, and Hearing Research, 41*, 374-384.

Roulstone, S., Loader, S., Northstone, K., Beveridge, M. and the ALSPAC team. (2002). The speech and language of children aged 25 months: Descriptive data from the Avon Longitudinal Study of Parents and Children. *Early Child Development and Care, 172*(3), 259-268.

Rust, J. (1996) *Weschler Objective Language Dimensions Manual*. London: The Psychological Corporation.

Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research, 36*, 105-140.

Shriberg, L. D., Friel-Patti, S., Flipsen, P. & Brown, R. L. (2000). Otitis media, fluctuant hearing loss, and speech-language outcomes: A preliminary structural equation model. *Journal of Speech, Language, and Hearing Research*, *43*, 100-120.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L. & Wilson, D. L. (1997). The speech disorders classification system (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research, 40,* 723-240.

Schwartz, R. G., Leonard, L. B., Folger, M. K. & Wilcox, M. J. (1980). Early phonological behaviour in normal-speaking and language disordered children: Evidence for a synergistic view of linguistic disorders. *Journal of Speech and Hearing Disorders, 45*, 357-377.

Selby, J. C., Robb, M. P. & Gilbert, H. R. (2000). Normal vowel articulations between 15 and 36 months of age. *Clinical Linguistics and Phonetics, 14*, 255-266.

Stewart, J. M., Hester, E. J. & Taylor, D. L. (1986). Prevalence of language, speech and hearing disorders in an urban preschool black population. *Journal of Childhood Communication Disorders*, *9,* 107–123.

Stoel-Gammon, C. (2007). Variability in speech acquisition. In S. McLeod, (Ed). *The International Guide to Speech Acquisition* (pp. 55-60) NY: Thomson Delmar Learning.

Tuomi, S. & Ivanoff, P. (1977). Incidence of speech and hearing disorders among kindergarten and grade 1 children. *Special Education in Canada*, *51,* 5–8.

Valanis, B. (1999). *Epidemiology in health care (3rd ed).* Stamford: Appleton & Lange.

Vogel Sosa, A. & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language*, *33,* 31-50.

Weiss, G. (1996). Research issues in longitudinal studies. In L. Hechtman (Ed). *Do they grow out of it? Long term outcomes of childhood disorders.* (pp. 1-16) Washington: American Psychiatric Press.

Williams, A. L. & Elbert, M. (2003). A prospective longitudinal study of phonological development in late talkers. *Language, Speech, and Hearing Services in Schools, 34*, 138-153.

Wren, Y., Miller, L. L., Roulstone, S., Peters T. J., & Emond, A. (in preparation) Prevalence and characteristics of children with speech impairment at age 8: A comparison of subtypes.

Appendix A. Object naming assessment at 2 years, adapted from Pagel Paden et al. (1987).

Target	Velar	Consonant cluster	Liquid	Fricative	Postvocalic singleton consonant
Brick	1	1	1		1
Brush		1	1	1	1
Clock	2	1	1		1
Cup	1				1
Fish				2	1
Flag	1	1	1	1	1
Flower		1	1	1	
Fork	1			1	1
Glasses	1	1	1	2	1
Light			1		1
Plate		1	1		1
Quack	2	1	1		1
Sock	1			1	1
Slide		1	1	1	1
Snake	1	1		1	1
Spoon		1		1	1
Potential occurrences	11	11	10	12	15

Although more opportunities are available for various target patterns only those noted here were scored.

Appendix B. Object naming assessment at 5 years, adapted from Pagel Paden et al. (1987).

Target	Velar	Consonant cluster	Liquid	Fricative	Postvocalic singleton consonant
Butterfly		1	1	1	
Calculator	2		2		1
Chips		1		2	
Clocks	2	2	1	1	
Glasses	1	1	1	2	1
Helicopter	1		1		1
Hippopotamus				1	1
Photograph	1	1	1	2	1
Present		2	1	1	
Pyjamas				2	1
Skirt	1	1		1	1
Spaghetti	1	1		1	
Squirrel	1	1	2	1	1
Starfish		1		3	1
String	1	1	1	1	1
Telephone			1	1	1
Television			1	1(ʒ only noted)	1
Three		1		1	
Toothbrush		1	1	2	2
Yellow			2		
Potential occurrences	11	15	16	24	14

Although more opportunities are available for various target patterns only those noted here were scored.

Box 1. Question to parent about the child's expressive language stage

Children learn to talk at different rates. Here are some stages of learning to talk. Tick the one which best describes your child.

Doesn't make any sounds

Makes babble noises that sound like talking or conversation

Uses single words and animal noises e.g., *miaow, gone, bikit* (biscuit)

Uses two words together e.g., *doggie gone, want dinner*

Uses 3-4 word sentences, e.g., *me want go out*

Figure 1. Derivation of participants for this study

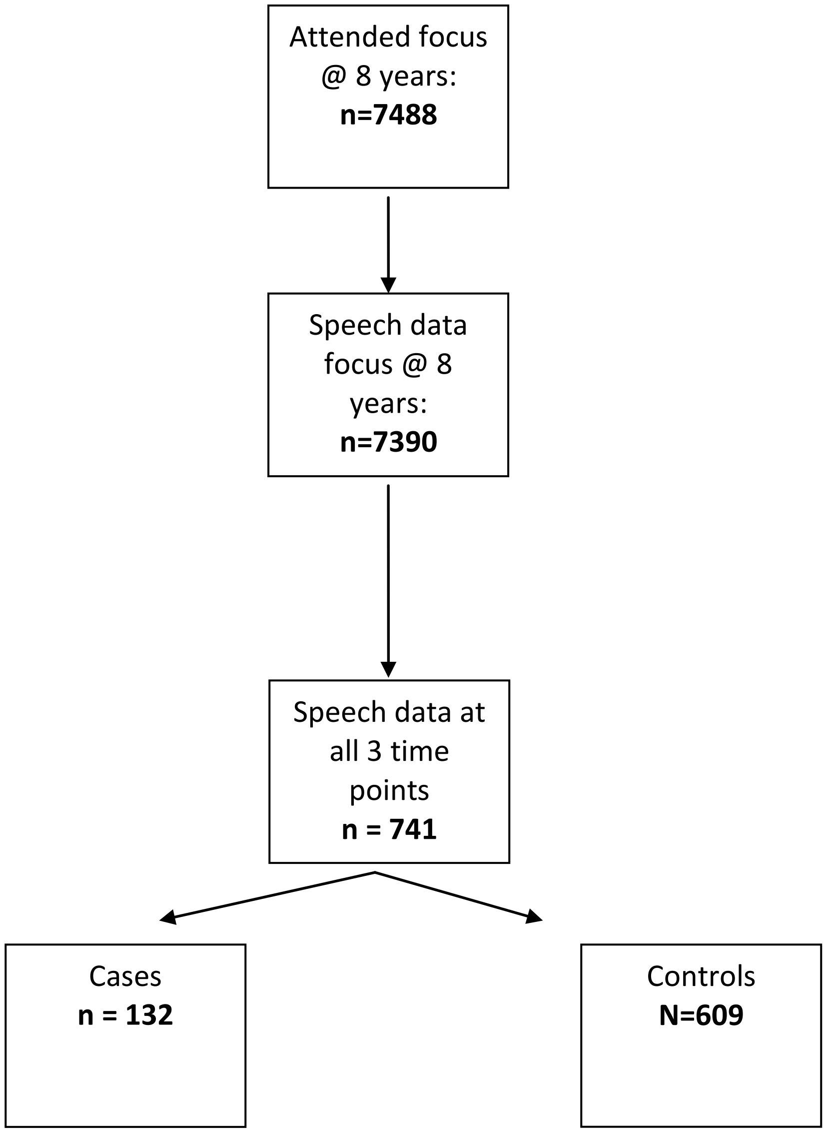

Table 1. Sample sizes available for analysis (and percentages)

	25 months		61 months	
	Cases	Controls	Cases	Controls
Postvocalic consonants	79 (16)	412 (84)	94 (17)	445 (83)
Velar consonants	68 (16)	365 (84)	94 (18)	440 (82)
Fricative consonants	71 (16)	380 (84)	94 (17)	445 (83)
Liquid consonants	58 (16)	316 (84)	94 (17)	444 (83)
Consonant clusters	68 (16)	359 (84)	94 (17)	445 (83)

Table 2. Gender, social class and age comparisons of cases and controls

		Cases	Controls
Gender	Male	79 (59.9)	312 (51.2)
	Female	53 (40.1)	297 (48.8)
Parental social class (lowest of parents)	Non-manual	71 (61.2)	344 (59.5)
	Manual	45 (38.8)	234 (40.5)
Maternal age (years)		29.4 (4.5)	29.6 (4.3)
Age at 2 years		108.2 (0.9)	108.2 (0.9)
Age at 5 years		268.9 (2.9)	268.4 (3.1)
Age at 8 years		103.1 (1.8)	103.4 (2.0)

For categorical variables n (%) are shown. For continuous variables mean (sd) are displayed.
For parental social class n=694 for all others it is 741.

Table 3. Mean proportion of errors and standard deviation

	2 years				5 years			
	Cases		Controls		Cases		Controls	
	n	Mean (sd)	n	Mean (sd)	n	Mean (sd)	n	Mean (sd)
Postvocalic consonants	79	24.7 (29.8)	412	18.4 (24.7)	94	7.8 (8.9)	445	5.9 (6.0)
Velar consonants	68	31.3 (27.9)	365	26.8 (25.8)	94	6.7 (13.2)	440	4.1 (9.1)
Fricative consonants	71	41.0 (30.0)	380	31.3 (25.0)	94	4.8 (12.8)	445	1.6 (4.3)
Liquid consonants	58	54.1 (29.5)	316	50.2 (30.1)	94	14.5 (16.8)	444	8.6 (13.8)
Consonant clusters	68	73.7 (25.4)	359	65.4 (29.6)	94	9.1 (15.6)	445	4.8 (8.8)

Table 4. Univariable Linear Regression

	2 years			5 years		
	n	Coef*	P-value	n	Coef*	P-value
Postvocalic consonants	491	6.3 (0.1, 12.5)	0.045	539	1.9 (0.5, 3.4)	0.010
Velar consonants	433	4.5 (-2.3, 11.3)	0.19	534	2.5 (0.3, 4.8)	0.025
Fricative consonants	451	9.7 (3.1, 16.2)	0.004	539	3.2 (1.7, 4.7)	<0.001
Liquid consonants	374	4.0 (-4.4, 12.4)	0.35	538	5.9 (2.7, 9.1)	<0.001
Consonant clusters	427	8.2 (0.7, 15.8)	0.032	539	4.2 (1.9, 6.5)	<0.001

* Calculated as the difference in mean percentage of errors between cases and controls. For example for fricative consonants at 25 months, the mean percentage of errors is 9.7 percentage points higher than controls (41.0 - 31.3 from table 3).

Table 5. Expressive language stage and speech sound errors at 25 months

			Single words		2 word together		3/4 word sentences	
	n	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	P-value
Postvocalic consonants	491	33	52.4 (32.5)	119	27.3 (27.1)	339	13.4 (20.9)	<0.001
Velar consonants	433	19	44.4 (28.0)	104	37.3 (27.5)	310	23.2 (24.3)	<0.001
Fricative consonants	451	24	56.1 (26.6)	104	44.4 (27.1)	323	27.4 (23.5)	<0.001
Liquid consonants	374	13	69.2 (18.0)	77	62.9 (23.3)	284	46.7 (30.8)	<0.001
Consonant clusters	427	17	83.2 (14.2)	97	79.9 (19.1)	313	61.8 (30.7)	<0.001

P-values are based on linear regression.

Table 6. Longest average sentence and speech sound errors at 61 months.

	n	Coefficient (95%CI)	P-value
Postvocalic consonants	539	-0.62 (-0.94, -0.30)	<0.001
Velar consonants	534	-0.45 (-0.66, -0.23)	<0.001
Fricative consonants	539	-0.35 (-0.67, -0.03)	0.033
Liquid consonants	538	0.03 (-0.12, 0.18)	0.70
Consonant clusters	539	-0.25 (-0.45, -0.04)	0.020

Linear regression was used to estimate the difference between the longest average sentences for a 10-percentage point increase in speech sound errors (e.g. as the postvocalic consonants error rate increases by 10 percentage points the sentence length decreases by 0.62 words).

Table 7. Multivariable logistic analyses

	25 months			61 months		
	n	OR	P-value	n	OR	P-value
Postvocalic consonants	491	1.06 (0.96, 1.17)	0.24	539	1.44 (1.03, 2.01)	0.033
Velar consonants	433	1.03 (0.57, 1.72)	0.52	534	1.21 (1.00, 1.47)	0.054
Fricative consonants	451	1.12 (1.01, 1.24)	0.030	539	1.75 (1.22, 2.51)	0.002
Liquid consonants	374	1.03 (0.93, 1.13)	0.62	538	1.24 (1.07, 1.43)	0.003
Consonant clusters	427	1.08 (0.97, 1.20)	0.15	539	1.32 (1.09, 1.58)	0.004

Adjusted for parental social class, maternal age, child's exact age at the relevant clinic and the 8-year clinic, expressive language (2 years only) and average longest sentence (5 years only).