Predicting Language Change Between 3 and 5 Years and Its Implications for Early Identification

WHAT’S KNOWN ON THIS SUBJECT: Early speech and language delays are risk factors for later developmental and social difficulties. It is easier to identify them retrospectively than prospectively. Population characteristics and prevalence rates make screening problematic.

WHAT THIS STUDY ADDS: Using data from a birth cohort, this study identifies predictors of language performance at 5 years and 4 patterns of change between 3 and 5 years, comparing those who change with those whose profile remains low across time points.

abstract

BACKGROUND AND OBJECTIVE: Early language delays across the preschool period have important implications for children, parents, and services raising the significance of early identification. Screening tests are an appealing solution but have proved problematic. A combined risk model would seem promising but has yet to be tested. The goal of this study was to examine the factors that predict language change in a nationally representative sample of children between 3 and 5 years when most children are identified as being in need of services.

METHODS: By using data from children (n = 13,016) in the Millennium Cohort Study (a national UK birth cohort), linear regression was used to predict 5-year performance from 3-year test performance data coupled with sociodemographic and within-child factors and indicators of parental concern. Patterns of change were identified and logistic regression was used to predict the difference between children for whom profiles change and those for whom they do not.

RESULTS: The final model (predicting 32% of the variance) included maternal education, pattern construction, behavior, language concerns, and 3-year vocabulary. Four change patterns were identified: one consistently low (n = 201), one consistently high (n = 12,066), a group that is resilient (n = 572), and one with a declining profile (n = 177). The models accurately predicted 71% of the declining group and 99% of the resilient group. Maternal education (odds ratio: 0.49) and behavior (odds ratio: 0.9) were significant predictors for the former and maternal education (odds ratio: 0.6) and pattern construction (odds ratio: 1.03) the latter.

CONCLUSIONS: Early identification of delayed language remains problematic but, once identified, there are key indicators that predict which children are likely to be more or less at risk across time. The implications are discussed in terms of policy and practice. Pediatrics 2012;130:e132–e137

AUTHORS: James Law, PhD,a Robert Rush, MSc,b Chantelle Anandan, MSc,c Marie Cox, BSc,d and Rachael Wood, MSc,e

aInstitute of Health and Society, School of Education, Communication and Language Sciences, Newcastle University, Newcastle-upon-Tyne, United Kingdom; bQueen Margaret University, Edinburgh, United Kingdom; cDepartment of General Practice and dCollege of Medicine, University of Edinburgh, Edinburgh, United Kingdom; and eInformation Services Division NHS National Services, Edinburgh, Scotland

KEY WORDS
child development, cohort studies, language screening—developmental

ABBREVIATIONS
BAS II—British Ability Scales Second Edition
CLL—consistently low language
IVL—increasingly vulnerable language
RL—resilient language
TL—typical language

www.pediatrics.org/cgi/doi/10.1542/peds.2011-1673
doi:10.1542/peds.2011-1673
Accepted for publication Mar 16, 2012
Address correspondence to James Law, PhD, Institute of Health and Society, School of Education, Communication and Language Sciences, Newcastle University, Newcastle-upon-Tyne, NE1 7RU United Kingdom. E-mail: james.law@ncl.ac.uk

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).
Copyright © 2012 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: The study was funded by the Chief Scientist Office, Scotland (HSR CSO17).
Although there is good evidence for the potential long-term consequences of early language delay,1,2 the process of early identification remains a live issue.3–7 Capturing children’s performance using screening tests is appealing as a concept but difficult to realize, the alternative being a hybrid model involving specific assessments of behavior and risk at differing time points.8,9 Combining performance with risk factors looks to be promising given what is known about factors associated with such delays.10–12 The role of parental concerns plays a key in this issue.13 The changing patterns of development over the early years are clearly central to the discussion, reflected in attempts to capture the difference between those children who experience consistent patterns of development and those who do not.14–16

In this study, we examined the factors that predict language change in a nationally representative sample of children between 3 and 5 years of age, a time when most children are identified as being in need of services. Two questions are addressed. To what extent is it possible to predict 5-year language performance from a range of sociodemographic and within-child predictors coupled with earlier language performance? Is it possible to characterize accurately a group whose performance changes over time relative to a group whose performance does not change?

METHODS

Data Sources

The Millennium Cohort Study17 is a national birth cohort of children born in the United Kingdom in 2000 and 2001. To date it includes four sweeps (9 months, 3, 5, and 7 years) of which the present analysis includes data collected on children at 3 and 5 years. Specifically the analysis includes all children who completed the naming vocabulary scale of the British Ability Scales (BAS II)19 at 3 and 5 years (n = 13,016; males: n = 6,566 [50.4%]; females: n = 6,450 [49.6%]). This is an expressive language test used between 2 and 8 years in which the child is required to name pictures of objects. We controlled for nonverbal performance by using the 3-year pattern construction subscale of the BAS II.

Variables

A range of relevant predictor variables was adopted. Mother’s education was used as a measure of sociodemographic status, split at above and below grade C in the national General Certificate of Secondary Education examination at 16 years of age. We identified whether children were “small for gestational age” if they were born after >259 gestational days weighing <2515 g and whether there had been any longstanding health concerns at 3 years. Gross motor and fine motor items were included from the Denver Developmental Screening Test reported at 9 months.20 In this context, “delay” was determined, following the authors’ recommendation, when an infant has not reached a milestone that 90% of infants have reached. Parental concerns about language development and hearing at 3 years were grouped into “some concerns” or “none/not applicable.” Behavior was assessed by using the parent report version of the Strengths and Difficulties Questionnaire.21

Analytic Strategy

The analysis was conducted in 2 phases. In the first, we used linear regression with BAS II naming vocabulary scale at 5 years as the dependent variable. In the second, language performance was categorized at 3 and 5 years, examining the productivity figures (sensitivity, specificity, and both positive and negative predictive ability) of the 3-year naming vocabulary scale in predicting 5-year vocabulary. Logistic regression was then used to examine factors associated with a series of change profiles. In each case, the candidate covariates were the same. Block 1 includes gender, maternal education, and whether the child was born small for gestational age. Block 2 adds the early identification of gross and fine motor difficulties at 10 months of age. Block 3 includes parental concerns about the child’s hearing and whether the child had experienced any long-term illness. Block 4 adds the pattern construction scale from the BAS II; block 5 adds behavior; and block 6 adds whether the parent expressed any concerns about language development when the child was seen at 3 years.

All analyses were performed by using SPSS version 17.0 for Windows (SPSS Inc, Chicago, IL). Data were weighted to the UK population.

Operationalizing the Groups

In the second phase of the analysis, the BAS II naming vocabulary outcome was split into 4 groups. Following convention,22 we defined children as being language delayed if they scored less than −1.5 SDs of the weighted sample distribution of age-related normed test scores on the BAS II naming vocabulary scale. The typical language group (n = 12,066) had scores within normal limits at both 3 and 5 years. An increasingly vulnerable language (IVL) group (n = 177) had typical development at 3 years but language delay by 5 years; a resilient language (RL) group (n = 572) was language delayed at 3 years but developing typically by 5 years; and a consistently low language (CLL) group (n = 201) had language delay at both time points. A McNemar test demonstrated that there was a significant change in scores on the BAS II from 3 to 5 years (χ² = 1554.965, df = 1, P < .001). Mean ± SD scores for the BAS II for the 4 groups are provided in Table 1.

Modeling Steps

A series of univariable and multivariable regression models were used to identify
variables that were independently associated with the outcome (BAS II at 5 years) in 2 phases. At each step, attention was paid to the effects of adjusting for other variables on the regression coefficients; confidence intervals and significance levels are reported. A series of univariable regression analyses were performed to test the strength of the association of each variable with the outcome variable. Those significant at $P < .1$ were retained in the next step of the analysis. To identify the best explanatory combination, thus increasing confidence in the robustness of the final models, any variables remaining significant at the 5% level within the multivariable analyses were carried forward to the final multivariable analysis.

**RESULTS**

For the first phase, a decrease was seen in scores across the models for boys, being small for gestational age, early gross and fine motor delays, long-standing health conditions, and hearing concerns; the changes started significant but dropped out once and were no longer statistically significant other blocks were added (Table 2).

In the final model (model 8), whether cohort members were born small for their due date was no longer significant and the magnitude of the independent effect of the other variables was reduced. The addition of the 3 years’ score again doubled the variance, to explain 32% of the variation in the BAS II at 5 years. Of this, half was accounted for by the initial BAS II scores and the remainder by the other factors, suggesting that

---

**TABLE 1** BAS II Naming Vocabulary Scale at 3 and 5 Years of Age (Mean and 95% confidence intervals)

<table>
<thead>
<tr>
<th>Group</th>
<th>BAS II Naming Vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL</td>
<td>52.52 (52.15, 52.5)</td>
</tr>
<tr>
<td>IVL</td>
<td>43.53 (42.65, 44.45)</td>
</tr>
<tr>
<td>RL</td>
<td>29.78 (29.4, 30.15)</td>
</tr>
<tr>
<td>CLL</td>
<td>27.15 (26.67, 27.62)</td>
</tr>
</tbody>
</table>

TL, typical language.

**TABLE 2** Factors Predicting 5-Year Language Performance

<table>
<thead>
<tr>
<th>Factor</th>
<th>Group 1 (n = 14,069)</th>
<th>Group 2 (n = 14,069)</th>
<th>Group 3 (n = 13,935)</th>
<th>Group 4 (n = 14,069)</th>
<th>Group 5 (n = 13,935)</th>
<th>Group 6 (n = 13,935)</th>
<th>Group 7 (n = 13,935)</th>
<th>Group 8 (n = 13,935)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material Education</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small for gestational age</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-standing health conditions</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hearing concerns</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language concerns</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS II vocabulary at 3 years</td>
<td>0.22 (0.21 to 0.23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Data are presented as $\beta$ (95% confidence interval).

---

LAW et al

Downloaded from by guest on June 1, 2017
enhancing child performance with relevant history data would improve a model of early identification.

In the cross-classification presentation (Table 3), we see the characteristic pattern\(^3\,4\) of high specificity (0.95) and low sensitivity (0.53), where prevalence rates are relatively low, suggesting that the 3-year score was very good at determining who was not likely to have a subsequent difficulty but less accurate in predicting who was. This finding is reflected both in the positive (0.26) and negative (0.96) predictive abilities. With a threshold of \(-1.5\) SDs, the prevalence rates were 5.9% at 3 years, dropping to 2.9% at 5 years; however, only those included at time 1 were included in time 2.

Four patterns of change were then plotted (Fig 1). Table 4 provides the final model for the difference between the IVL and CLL groups.

Univariable associations indicated that only maternal education, gross motor skills, and behavior significantly distinguished between the 2 groups. Maternal educational level was consistently strongly associated across all the blocks, with those in the poor maternal education group being half as likely to be in the IVL group. When controlling for maternal education, behavior remained associated, with those with greater behavior difficulties being more likely to be in the CLL group, whereas gross motor skills dropped out of the analysis. The final model correctly predicted 57% of those in the CLL group and 71% of those whose performance declines (overall: 64%).

Table 5 reports comparable results for the RL and CLL groups. Univariable associations found that maternal education, gross motor skills, pattern construction, and behavior were significant \((P = .1)\). For the multivariable models, maternal education was again significantly differentiating the 2 groups, with the CLL group having mothers with the lowest educational attainment. Gross motor skills were not significant in the presence of maternal education. In model 3, pattern construction remained in the final model while behavior dropped. The capacity of these models to predict at an individual level in terms of percent correctly identified is generally reasonable although the patterns can sometimes appear anomalous. This was the case for the CLL group, of whom just 2% were correctly classified relative to the RL group, of whom virtually all were correctly classified (Table 5). It is important to stress that “the classification table is most appropriate where classification is the stated goal of the analysis; otherwise it should only supplement more rigorous methods of assessment of fit.”\(^2\,3\)

**DISCUSSION**

Our results suggest that different characteristics distinguish the increasingly vulnerable and resilient groups. In both cases, although those least likely to change had mothers with lower levels of education, the CLL group had a higher level of behavior difficulties than those whose scores subsequently drop. By contrast, it is the child’s nonverbal performance that seems to trigger resilience. Both highlight the need to include supplementary within-child and environmental information when weighing risk. It seems easier to identify reduced risk in those with initial delays but less easy to identify those whose skills are likely to decrease. The findings demonstrate well the inherent problem with developmental screening. Even with a
TABLE 4 Difference Between the CLL Group (n = 201) and the IVL Group (n = 177)  

<table>
<thead>
<tr>
<th>Factor</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal education</td>
<td>0.450*** (0.298 to 0.680)</td>
<td>0.452*** (0.298 to 0.685)</td>
<td>0.480** (0.302 to 0.795)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross motor delay (Denver)</td>
<td>—</td>
<td>0.485 (0.235 to 1.013)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior (SDQ)</td>
<td>—</td>
<td>—</td>
<td>0.923*** (0.885 to 0.962)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>376</td>
<td>375</td>
<td>289</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLL group correct %</td>
<td>60.5</td>
<td>64.8</td>
<td>56.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IVL group correct %</td>
<td>59.2</td>
<td>55.0</td>
<td>70.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall correct %</td>
<td>59.9</td>
<td>60.2</td>
<td>64.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are presented as odds ratio (95% confidence interval), n, or %. Denver, Denver Developmental Screening Test; SDQ, Strengths and Difficulties Questionnaire.  
***P < .001, **P < .01, *P < .05.

The strength of the study is that it is derived from a nationally representative sample by using a well-standardized measure of expressive vocabulary and, perhaps most significantly, a measure that is repeated across the vitally important 2-year period before school entry. We carried out an analysis of different thresholds (−1 and −2 SDs) but this proved almost identical to those in the present analysis, adding confidence to the interpretation and obviating the problem of marginal test differences between groups. It is also important to acknowledge that although many of the associations with independent factors in the logistic regression models are statistically significant, it does not mean that the resultant model necessarily classifies the groups well on the dependent variable.23

Implications for Policy and Practice

Policy makers and practitioners are faced with a dilemma. On the one hand, there is strong evidence that early language delays make a child vulnerable to later difficulties. This relationship becomes clearer as the young child moves toward compulsory schooling.24 On the other hand, results from this and many other studies suggest that the adoption of formal procedures, even measures as psychometrically robust as the BAS II, remains problematic. The data here suggest that the key is to combine child performance data, including earlier performance and both verbal and nonverbal concurrent performance, together with aspects of sociodemographic risk. But even this would be difficult if only carried out at a single time point, forcing us to conclude that there remains a strong case for the population monitoring of risk groups before the school years, whether conducted by pediatricians in the medical home or health visitors in community health center and that delayed language acquisition is probably one of the most useful and accessible litmus tests of early childhood well-being. Such a conclusion emphasizes the potential role for “proportionate universalism,”25 by which universal services such as screening and early identification are tailored to the level of disadvantage, not restricted to those with the highest levels of need. Our findings suggest that patterns of change in early language skills are potentially one method of differentiating such needs.

Although attention has been paid in the present analysis to significant within-child and environmental factors, we have of course overlooked one of the most important potential environmental modifications to which children of this age are exposed; namely, early years of education and specifically speech and language therapy services.26 The latter have been shown to be effective with the type of problems that would be identified by a naming vocabulary measure, but we unfortunately have no way of knowing which children received such services.27

Implications for Research

Findings from the present analysis would warrant replication in comparable large-scale representative data sets. Better understanding of the relationship between patterns of change and the potential implications of intervention

TABLE 5 Difference Between the CLL Group (n = 201) and the RL Group (n = 572)  

<table>
<thead>
<tr>
<th>Factor</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal education</td>
<td>0.481*** (0.346 to 0.668)</td>
<td>0.486*** (0.349 to 0.675)</td>
<td>0.547*** (0.385 to 0.766)</td>
<td>0.526** (0.355 to 0.78)</td>
<td>0.596* (0.396 to 0.896)</td>
</tr>
<tr>
<td>Gross motor delay (Denver)</td>
<td>—</td>
<td>0.672 (0.402 to 1.12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pattern construction, (BAS II)</td>
<td>—</td>
<td>—</td>
<td>1.037*** (1.021 to 1.052)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior, (SDQ)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.989 (0.94 to 1)</td>
<td>0.976 (0.945 to 1.007)</td>
</tr>
<tr>
<td>n</td>
<td>770</td>
<td>770</td>
<td>761</td>
<td>581</td>
<td>575</td>
</tr>
<tr>
<td>CLL group correct %</td>
<td>0</td>
<td>0</td>
<td>7.4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>RL group correct %</td>
<td>100</td>
<td>100</td>
<td>97.6</td>
<td>100</td>
<td>99.6</td>
</tr>
<tr>
<td>Overall correct %</td>
<td>74.1</td>
<td>74.1</td>
<td>75</td>
<td>76.4</td>
<td>77.5</td>
</tr>
</tbody>
</table>

Data are presented as odds ratio (95% confidence interval), n, or %. Denver, Denver Developmental Screening Test; SDQ, Strengths and Difficulties Questionnaire.  
***P < .001, **P < .01, *P < .05
would be helpful. Finally, it would be useful to explore the patterns of change at both a group and an individual level as children move through school to put figures on the burden of risk associated with early language difficulties.

CONCLUSIONS

This study is one of a small number that have used large-scale population data to examine patterns of language change across time but the first, to the best of our knowledge, to do so in the context of early identification. There is evidence to suggest that early language delays point to later difficulties; however, this finding does not readily translate into identification at an individual level. Although it is possible to anticipate language performance at school entry from a relatively predictable set of factors, it would seem that, beyond maternal education, those that predict risk and resilience are not necessarily the same. We could conclude that testing the language skills of 3-year-olds is a relatively good way of establishing which children will not have later difficulties, but there remains a need both to monitor poor performers and to improve our understanding of the mechanisms by which children’s language skills falter after seeming to develop normally.

REFERENCES

Predicting Language Change Between 3 and 5 Years and Its Implications for Early Identification

James Law, Robert Rush, Chantelle Anandan, Marie Cox and Rachael Wood

*Pediatrics*; originally published online June 11, 2012;
DOI: 10.1542/peds.2011-1673

Updated Information & Services
including high resolution figures, can be found at:
/content/early/2012/06/06/peds.2011-1673

Citations
This article has been cited by 2 HighWire-hosted articles:
/content/early/2012/06/06/peds.2011-1673#related-urls

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
/site/misc/Permissions.xhtml

Reprints
Information about ordering reprints can be found online:
/site/misc/reprints.xhtml

PEDIATRICS is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. PEDIATRICS is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 2012 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 0031-4005. Online ISSN: 1098-4275.
Predicting Language Change Between 3 and 5 Years and Its Implications for Early Identification
James Law, Robert Rush, Chantelle Anandan, Marie Cox and Rachael Wood
Pediatrics; originally published online June 11, 2012;
DOI: 10.1542/peds.2011-1673

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/early/2012/06/06/peds.2011-1673