

Genetic and phenotypic effects of phonological short-term memory and grammatical morphology in specific language impairment

M. Falcaro[†], A. Pickles^{*†}, D. F. Newbury[‡],
L. Addis[‡], E. Banfield[‡], S. E. Fisher[‡], A. P.
Monaco[‡], Z. Simkin[†], G. Conti-Ramsden[†]
and The SLI Consortium¹

[†]Faculty of Medical and Human Sciences, The University of Manchester, Manchester, and [‡]The Wellcome Trust Centre for Human Genetics, Oxford, United Kingdom

*Corresponding author: A. Pickles, Biostatistics Group, The University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK. E-mail: andrew.pickles@manchester.ac.uk

Deficits in phonological short-term memory and aspects of verb grammar morphology have been proposed as phenotypic markers of specific language impairment (SLI) with the suggestion that these traits are likely to be under different genetic influences. This investigation in 300 first-degree relatives of 93 probands with SLI examined familial aggregation and genetic linkage of two measures thought to index these two traits, non-

word repetition and tense marking. In particular, the involvement of chromosomes 16q and 19q was examined as previous studies found these two regions to be related to SLI. Results showed a strong association between relatives' and probands' scores on non-word repetition. In contrast, no association was found for tense marking when examined as a continuous measure. However, significant familial aggregation was found when tense marking was treated as a binary measure with a cut-off point of -1.5 SD, suggestive of the possibility that qualitative distinctions in the trait may be familial while quantitative variability may be more a consequence of non-familial factors. Linkage analyses supported previous findings of the SLI Consortium of linkage to chromosome 16q for phonological short-term memory and to chromosome 19q for expressive language. In addition, we report new findings that relate to the past tense phenotype. For the continuous measure, linkage was found on both chromosomes, but evidence was stronger on chromosome 19. For the binary measure, linkage was observed on chromosome 19 but not on chromosome 16.

Keywords: Familial aggregation, grammatical morphology, linkage analysis, phonological short-term memory, selected sample, specific language impairment

Received 19 June 2007, revised 6 September 2007, accepted for publication 25 September 2007

¹ Members of the SLI Consortium: The Wellcome Trust Centre for Human Genetics, University of Oxford: D. F. Newbury, E. Banfield, L. Addis, J. D. Cleak, S. E. Fisher, L. R. Cardon and A. P. Monaco. Cambridge Language and Speech Project (CLASP): M. J. Merricks and I. M. Goodyer. Child and Adolescent Psychiatry Department and Medical Research Council Centre for Social, Developmental, and Genetic Psychiatry, Institute of Psychiatry: E. Simonoff and P. F. Bolton. Newcomen Centre, Guy's Hospital: V. Slonims and G. Baird. Department of Child Health, University of Aberdeen: A. Everitt, E. Hennessy, D. Shaw and P. J. Helms. The Raeden Centre and Grampian University Hospitals Trust: A. D. Kindley. Speech and Hearing Sciences, Queen Margaret University College: A. Clark and J. Watson. Department of Reproductive and Developmental Sciences, University of Edinburgh: A. O'Hare. Molecular Medicine Centre, University of Edinburgh: J. Seckl. Department of Speech and Language Therapy, Royal Hospital for Sick Children, Edinburgh: H. Cowie. Department of Educational and Professional Studies, University of Strathclyde: W. Cohen. Academic Unit of Neurology, University of Sheffield Medical School: J. Nasir. Department of Experimental Psychology, University of Oxford: D. V. M. Bishop. Human Communication and Deafness, School of Psychological Sciences, University of Manchester: Z. Simkin and G. Conti-Ramsden. Biostatistics Group, University of Manchester: M. Falcaro and A. Pickles.

Specific language impairment (SLI) is a heterogeneous disorder, presenting with a variety of profiles, all of which involve difficulties with language learning in the absence of possible explanatory factors such as low non-verbal IQ, hearing impairment or neurological damage.

Findings of a strong familial aggregation in SLI (Neils & Aram 1986; Tallal *et al.* 1989) have spurred a number of studies examining more closely the potential genetic contribution to SLI. Nonetheless, there is still great uncertainty as to which traits or elements of the SLI profile are key and whether these phenotypic traits are under similar or different genetic influence.

Bishop *et al.* (2006) examined both non-word repetition and verb grammatical morphology in twin pairs selected for risk of language impairment. The results of their bivariate twin analysis suggested that both non-word repetition and verb grammatical morphology were significantly heritable and discriminated children at risk of language problems from those with low risk. Interestingly, and in line with previous

findings (Conti-Ramsden *et al.* 2001), there was little phenotypic overlap between the two deficits, but, in addition, Bishop *et al.* found the additive genetic variance for the two phenotypes to be largely distinct. Thus, it may not be the case that SLI is either a problem with phonological short-term memory capacity or a problem with a linguistic mechanism involved in grammatical morphology, but that within the SLI population there are children with language difficulties resulting from either or both sets of deficits, each with a distinct genetic aetiological contribution.

Genome-wide linkage screens of quantitative measures of language have been undertaken by The SLI Consortium (2002, 2004) using two samples of 98 (wave 1) and 86 (wave 2) families where at least one sibling met the SLI criteria. The two waves combined showed very strong linkage to chromosome 16 for non-word repetition (*SLI1* MIM606711). Linkage was also found to chromosome 19 (*SLI2* MIM606712), but the interpretation of the results at this location was more complex because the separate waves of families were primarily linked to two different phenotypes: Clinical Evaluation of Language Fundamentals (CELF-R) expressive language for wave 1 and non-word repetition for wave 2. In a third paper by the SLI Consortium (SLIC) (Monaco & The SLI Consortium 2007), multivariate techniques were used for the joint analysis of literacy and language measures in an attempt to identify common and specific effects at the *SLI1* and *SLI2* loci and to clarify the discrepancies between the profiles of phenotypes linked in each wave. In these SLIC studies, however, no measure of grammatical morphology was available.

The principal aims of this paper were to investigate the findings of Bishop *et al.* (2006) by evaluating the familial aggregation of non-word repetition and a tense-marking measure of grammatical morphology in first-degree relatives, to replicate previous linkage findings for non-word repetition and expressive language in the *SLI1* and *SLI2* regions in an entirely independent sample and to extend these findings by exploring possible specificity of linkage for grammatical morphology.

The study

Participants

Participants were originally part of the Manchester Language Study (Conti-Ramsden & Botting 1999a,b; Conti-Ramsden *et al.* 1997). Proband was recruited from 118 language units attached to English mainstream schools. All language units catering for primary school year 2 children were contacted. Centres enrolling children with global delay or hearing impairments were excluded, and two centres declined to participate. Approximately half of the children attending a language unit for at least 50% of the week were randomly sampled, yielding 242 children (185 males and 57 females) aged between 6 years and 2 months and 7 years and 10 months. These children were then reassessed at approximately 8 years and 11 years of age and also at 14 years of age when assessment was extended to first-degree relatives. Of the 124 (51%) families who agreed to take part in the study, 11 were not assessed because of alterations in family circumstances. From the 113 fully consenting and assessed families, 93 were selected for participation in the present study based on the following proband criteria:

- 1 Performance IQ of 80 or more and a minimum of one concurrent standardized language test score that fell at least 1 SD below the population mean at one of the longitudinal assessment stages.

- 2 No sensory neural hearing loss.
- 3 English as a first language.
- 4 No record of a medical condition likely to affect language.
- 5 No record of a co-morbid diagnosis of autism.

The sample considered in this paper consists therefore of 93 probands (68 males and 25 females) and their first-degree relatives, where the minimum age for participation was 6 years. Proband had a mean age of 14 years and 5 months (age range: 13 years and 1 months to 16 years and 2 months). There were 300 first-degree relatives: 93 fathers, 93 mothers, 35 male/24 female siblings over the age of 16 years and 26 male/29 female siblings between the ages of 6 and 16 years. The mean age was 44 years and 1 month for parents, 18 years and 8 months for older siblings and 12 years and 4 months for younger siblings.

Phenotypic measures

Proband was tested on a battery of language, literacy and general cognitive measures. Psychometric tests and interviews were then carried out on all consenting first-degree relatives. We focus on two phenotypic measures: the non-word repetition (NWR) test and the past tense (PT) task, which were available for both probands and relatives.

NWR test

The NWR is a test of non-word repetition designed by Gathercole & Baddeley (1996) as an instrument to assess phonological short-term memory. This test consists of 28 non-words of various length and complexity. There are seven each of two-syllable non-words (e.g. 'brufid'), three-syllable non-words (e.g. 'shimitet'), four-syllable non-words (e.g. 'malpirony') and five-syllable non-words (e.g. 'dexiptecastic'). The non-words are presented in random order and given using live voice with lips shielded to prevent lip-reading, rather than using a tape recording. This has been found to be preferable when working with distractible or young children (Adams & Gathercole 1995, 1996) as the examiner could more easily ensure that the participant's attention was engaged before presenting an item. The test is scored online with each of the 28 items judged as correctly or incorrectly repeated. Standardized scores were calculated using normative data for a British population aged 5 years to undergraduate obtained from S. E. Gathercole & D. V. M. Bishop (personal communications).

PT task

The PT task is a 52-item test of grammatical morphology developed by Marchman *et al.* (1999). Participants are shown a drawing of everyday activities and asked to verbally fill in the missing word that the assessor leaves out while reading a sentence to them. The task comprises both regular and irregular verbs and is of the following type: 'The boy is walking. He walks everyday. Yesterday, he ...'. Each answer is classified as correct or incorrect and the test scored as the total number of correct PT inflections. These raw scores are not standardized for age, and most participants over the age of 12 years are expected to reach the 'ceiling' of 52 correct items.

It has been suggested that, in particular among adults, tense marking may not be a trait skill measurable on a continuous dimension but instead may be a skill that is simply acquired or not acquired (Bishop 2005). We will consider both possibilities, that is PT is either a binary or a continuous measure with a ceiling.

Expressive language

Expressive language scores (ELS) were obtained from the CELF-R (Semel *et al.* 1987) for probands and siblings between 6 and 16 years of age.

Analysis

Standardization of the PT raw scores

While standardized scores were available for the NWR test, analyses of the PT task were more complex requiring us to first convert the raw scores into standard scores while accounting for the frequent test ceiling score of 52. Normative data for the standardization of the PT

raw scores were retrieved from three studies on SLI: 31 children aged between 6 and 12 years from Marchman *et al.* (1999), 100 11-year-old children recruited by Simkin & Conti-Ramsden (2001) and 64 subjects aged between 13 and 16 years assessed for the purpose of this study. These 195 general population participants together with the Manchester sample (relatives and probands with allowance for a mean difference of each from the general population) were used to remove the age trends and to derive estimates of the quantities (mean and SD) necessary for the standardization. Standard linear regression models deliver biased results when the dependent variable is subject to a ceiling effect (right censoring). In addition, there was substantial evidence for skill acquisition to be age limited, with scores not increasing with age beyond the early teenage years. A Tobit model was used to correct for this test ceiling (Tobin 1958) consisting of simultaneously fitting (1) a linear model for the non-censored observations (i.e. $y_i < 52$) and (2) a probit model for the censored observations (i.e. $y_i \geq 52$ and we observed 52) where the underlying latent variable for individual i y_i (i.e. the variable we would have observed if there had been no ceiling) was modelled as

$$y_i = \alpha_0 \times \text{group0} + \alpha_1 \times \text{group1} + \alpha_2 \times \text{group2} + \alpha_3 \times \text{age14} + \varepsilon_i$$

This allows (1) different group means (group0 = 1 for normative data, group1 = 1 for relatives of SLI probands and group2 = 1 for probands), (2) an age trend up to 14 years (the covariate age14 is equal to the subject's age if the individual is younger than 14 years and is fixed at 14 years otherwise) and (3) different variability (heteroscedasticity) by group and age by allowing the error term ε_i to be normally distributed with mean zero and variance σ_ε^2 given by

$$\sigma_\varepsilon = \exp\{s_0 \times \text{group0} + s_1 \times \text{group1} + s_2 \times \text{group2} + s_3 \times \text{age14}\}$$

where s_0 , s_1 , s_2 and s_3 are parameters and the exponential function is used to guarantee σ_ε to be a positive quantity. This model was estimated in the procedure gllamm (Rabe-Hesketh *et al.* 2002) and then used to calculate estimated age-specific means and standard deviations for the normal population:

$$\text{Mean(normal population)} : \hat{\mu}_{np} = \hat{\alpha}_0 + \hat{\alpha}_3 \times \text{age14}$$

$$\text{SD(normal population)} : \hat{\sigma}_{np} = \exp(\hat{s}_0 + \hat{s}_3 \times \text{age14})$$

with the standardized scores then being computed as $(y_i - \hat{\mu}_{np})/\hat{\sigma}_{np}$. The ceiling of the test was accordingly standardized as $(52 - \hat{\mu}_{np})/\hat{\sigma}_{np}$. These transformed scores will be denoted as stdPT in what follows.

Analysis of familial aggregation

To test for familial aggregation of PT and NWR scores in relatives of SLI probands, we fitted regression models (a Tobit model for stdPT and a linear regression for NWR) incorporating the corresponding proband's score as a covariate and controlling for proband's and relative's sex. For the PT task, we also considered familiarity of a binary measure of affectedness by using a cut-off of 1.5 SD below the mean and a logistic regression. Lack of independence between observations was accounted for by the use of robust/sandwich standard errors (Huber 1967).

Differences in relatives' mean score profiles were then inspected by grouping the relatives by their own score, and their proband's score, at the -1.5 SD cut-off on each of stdPT and NWR.

Genotypic data

Participants were asked to provide a buccal swab sample for DNA analysis. Families with no full siblings ($n = 27$) or only non-consenting siblings ($n = 16$) or where the proband had co-occurring medical illness such as autism ($n = 8$), had non-Caucasian ancestry ($n = 1$) or contributed to a previous SLIC collection ($n = 1$) were excluded. This left 40 families for genotyping.

DNA was extracted using standard protocols and quantified with Pico Green (Molecular Probes, Eugene, OR, USA). If necessary,

samples were pre-amplified using Genomiphi DNA amplification kits (GE Healthcare, Amersham, UK). Eight microsatellite markers were amplified across a 10.5-Mb region of chromosome 16q (16q23.1–16q24.2) and nine microsatellite markers from a 23.5-Mb region of chromosome 19 (19q12–19q13.42). DNA was amplified using fluorescently labelled primers (Applied Biosystems, MWG Biotech, Foster City, CA, USA). Polymerase chain reaction products were pooled allowing concurrent detection by ABI 3700 sequencers (Applied Biosystems). Data were extracted using GENESCAN (version 3.1) and GENOTYPER (version 2.0) software (Applied Biosystems), manually verified and checked for inconsistencies within GENETIC ANALYSIS software (version 2, A. Young). Marker haplotypes were generated within GENEHUNTER 2.0 (Kruglyak *et al.* 1996), and all chromosomes showing an excessive number of recombinations were re-examined at the genotype level. The Integrated Genotyping System (R. Mott) was used for the storage of genotypic data. Sex-averaged maps were taken from the deCODE (Iceland) company's genetic map (Kong *et al.* 2002) supplemented with data from the human genome map (University of California, Santa Cruz). Markers were selected on average one every 10 cM.

Genotype, phenotype and map data were uploaded into GENEHUNTER 2.0 and used to calculate the likelihood of sharing 0, 1 or 2 alleles for each possible sibling pair at increments of 1 cM. These multipoint identity-by-descent (IBD) values were then directly used for linkage analyses as described below.

Linkage analysis

The sibships used for linkage analysis were not a random sample but were ascertained through the proband children of the Manchester Language Study. Selected samples are known to provide higher power to detect quantitative trait loci (QTL) effects (Carey & Williamson 1991), but they require methods that adjust for, or are robust to, the sample selection.

A simple and convenient way to analyse selected samples consists of using a DeFries–Fulker (DF) model (Fulker *et al.* 1991), which is a regression-based method based on the idea that the sibling's phenotypic score tends to regress back towards the mean of the general population as a function of shared environmental effects and the proportion of alleles shared IBD with the proband. Letting i be an index for sibship and j distinguish between siblings in the same sibship, a basic DF model consists in coupling each sibling with his/her corresponding proband and fitting the linear regression model:

$$C_{ij} = \beta_0 + \beta_1 P_i + \beta_2 \pi_{ij} + \varepsilon_{ij}$$

where C_{ij} represents the sib's phenotypic score; π_{ij} , the estimated proportion of alleles shared IBD at a given location along the chromosome; P_i , the proband's phenotypic score and β_0 , β_1 and β_2 are parameters. The error term ε_{ij} is assumed to be normally distributed with mean 0 and variance σ^2 . Testing $\beta_2 = 0$ corresponds therefore to test for linkage, and the corresponding t -statistics can be used to compute approximate LOD scores: $\text{LOD} = t^2/(2 \ln 10)$. The direction of the alternative hypothesis depends on which tail of the distribution the probands were selected from and on the scaling of the data. In our case, the t s of interest are those having negative values, calling for a one-sided test: $\beta_2 = 0$ vs. $\beta_2 < 0$.

To account for the right censoring arising from the ceiling for the stdPT scores, we used a Tobit DF model instead of a standard linear regression model. To avoid inflation of the LOD scores because of residual within-family correlation among our multiple sibpair families, the LOD scores were calculated from t -statistics based on the cluster robust parameter covariance matrix (Huber 1967).

Although the probands in our sample were likely to have poor language and literacy skills, ascertainment was not specifically based on the phenotypes under study and so not all the probands had low scores on PT and NWR. For each phenotype, we imposed a proband criterion such that only those scoring below a certain cut-off were considered probands. Following previous practice (Gayán *et al.* 1999), linkage analyses were repeated for a range of proband phenotypic cut-offs, with scores more than 1, 1.5, 1.75 and 2 SDs below the mean of the normal population.

To compare results with those of The SLI Consortium (2002, 2004) studies, we also applied the conventional Haseman–Elston (HE)

(Haseman & Elston 1972) linear regression of the squared difference in quantitative trait values of sibpairs on $\hat{\pi}$. This method is relatively robust to deviations from normality and to non-random sampling (Elston & Cordell 2001), but it requires a large number of siblings in order to have a reasonable power to detect linkage (Amos *et al.* 1989; Blackwelder & Elston 1982). A standard HE model with robust standard errors was fitted separately to the NWR and ELS phenotypes. For the continuous PT phenotype, this model could not be fitted because of the presence of censored observations. However, when considering PT as a binary measure (above/below the -1.5 SD cut-off), a logistic version of the HE model was used.

As the phenotypes were correlated and the genetic markers were close to each other, a classical Bonferroni correction of the linkage P values would be too conservative. We therefore preferred, as suggested by Elston (1998), to report single-location P values or LOD scores rather than those adjusted for multiple comparisons (Grigorenko *et al.* 2000; Knopik *et al.* 2002). For the peaks of the LOD scores, we also computed empirical P values using Monte Carlo permutation testing with 100 000 replicates. The IBD probabilities were randomly shuffled through a two-step procedure consisting of permutations across sibships of the same size and among pairs within each sibship (Shete *et al.* 2003). This procedure allows for dependent sibpairs and preserves the correlation structure of the data.

All the statistical analyses in this paper were performed within STATA, version 9 (StataCorp 2005).

Results

Descriptive statistics and standardization of PT

Descriptive statistics for PT, stdPT and NWR are reported in Table 1. As expected, on average, probands performed worse than their relatives on both tests. Six of 84 (7%) probands but 70 of 204 (34%) relatives achieved the maximum score of 52 on the PT task. The Spearman's rank correlation between NWR and PT was 0.14 for relatives and 0.35 for probands.

Figure 1 shows the estimated normal population's age-specific mean and 1 SD interval around the mean for the PT task along with the scatter plot of the raw scores from the three participant groups included in the standardization.

Past tense and NWR did not perform equally as markers of SLI. The percentage of subjects scoring below -1.5 SD on stdPT was 21% among relatives and 69% among probands, while for NWR, the percentages below -1.5 SD were 8% for

Table 1: Descriptive statistics for PT (raw scores), stdPT (standardized PT scores) and NWR (standard scores)

Variable	N (censored obs)	Mean*	SD*
PT (raw scores)			
Relatives	204 (70)	50.64	8.49
Probands	84 (6)	43.09	8.24
stdPT (standardized scores)			
Relatives	200 (68)	-0.28	2.85
Probands	84 (6)	-3.45	3.37
NWR (standard scores)			
Relatives	218	99.26	14.82
Probands	91	86.96	19.59

*For PT and stdPT, the mean and variance were estimated through interval regression to account for the censoring.

relatives and 41% for probands. Thirty-two (38%) probands performed below -1.5 SD on both measures and 21 (25%) performed above -1.5 SD on both measures. Twenty-six (31%) probands scored poorly on the stdPT alone, while only five (6%) scored poorly on NWR alone.

Analysis of familial aggregation

NWR possessed a clear familial component: probands' score on NWR was found to be strongly associated with relatives' low NWR scores ($P < 0.001$). By contrast, we found no significant association ($P = 0.6$) between the relatives' and their probands' stdPT scores. This result was unchanged when we included over-regularization errors (i.e. incorrect attempts at marking, e.g. thrown for threw) as correct responses. However, when using the binary measure of the stdPT score with the -1.5 SD proband cut-off, a significant familial aggregation was found: the odds of being affected for relatives of affected probands were around 2.5 times higher than those with an unaffected proband (OR = 2.47, 95% CI: 1-6.08, $P = 0.05$). A marked deficit in this skill may therefore be familial.

Linkage analysis

An assessment of familiarity gives an indication of the overall extent of likely familial variation but considerations of statistical power and aetiological complexity mean that its absence does not exclude the possibility of specific genetic effects.

After excluding families with missing or insufficient genetic and phenotypic data, 33 and 32 families contributed to the linkage analyses on chromosomes 16 and 19, respectively. DeFries-Fulker models were fitted to available proband-sibpairs (56 for chromosome 16 and 53 for chromosome 19). For the ELS phenotype, for which results are presented merely for comparison, fewer pairs were available (24 for chromosome 16 and 23 for chromosome 19) because the test was administered only to probands and child siblings (ages 6-16 years). As the ascertainment of the probands was not specifically based on the phenotypic measures under study, DF model-based QTL analyses were repeated imposing several proband selection criteria.

The LOD score profiles for chromosomes 16 and 19 for different proband selection cut-offs are displayed in Fig. 2 and the corresponding max LOD scores are displayed in Table 2.

For comparison, standard HE models were fitted to NWR and ELS and a logistic HE model was carried out for the binary PT measure. Unlike the DF model, the HE method involved all the possible sibpairs and not only those formed by a proband and co-sib (chromosome 16: 72 for PT, 80 for NWR and 28 for ELS; chromosome 19: 67 for PT, 75 for NWR and 27 for ELS). The LOD scores obtained from HE models are displayed in Fig. 3.

Table 3 summarizes the peaks of maximal linkage and their relative position as identified by the DF and HE models.

The LOD score profiles from the two models show similar patterns. It is, however, to be noticed that many of the LOD scores are artificially inflated because the empirical P values yield much reduced significance compared with the nominal P values that one would calculate for LODs of this magnitude using asymptotic considerations.

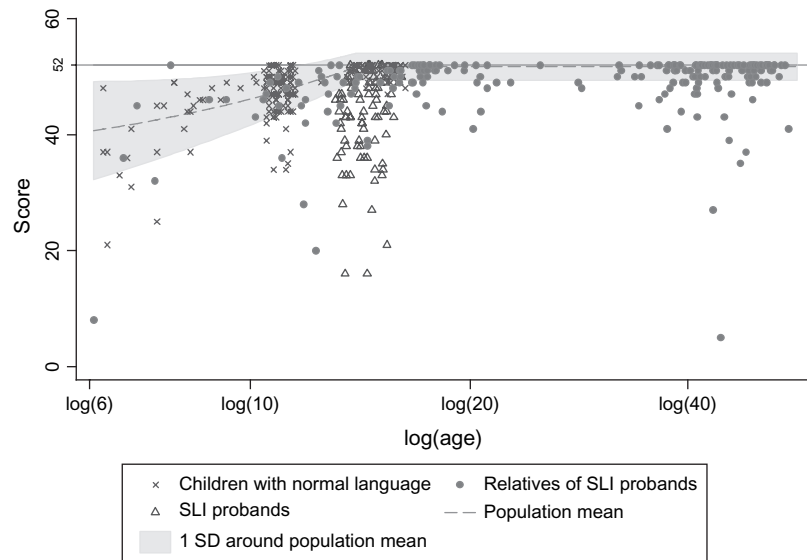


Figure 1: Estimated mean and ± 1 SD around the mean of the normal population for the PT task.

For the NWR phenotype, evidence of linkage was found on chromosome 16 by both DF (max LOD = 1.69, empirical $P = 0.015$) and HE models (max LOD = 1.54, empirical $P = 0.002$); evidence for linkage on chromosome 19 was somewhat weaker (empirical $P = 0.03$ and 0.07 by using, respectively, DF and HE models).

DeFries–Fulker linkage analysis gave strong evidence (empirical $P = 0.007$) for the existence of a QTL influencing expressive skills on chromosome 19. The HE method identified the same region but with reduced significance.

For the PT phenotype, some additional caution in interpreting the linkage analysis results was needed. When PT was measured on the continuous scale, the DF models gave max LOD score of 1.80 on chromosome 16 and of 2.20 on chromosome 19 (empirical $P = 0.01$ and 0.006, respectively). However, for the binary PT measure, linkage was only observed on chromosome 19 (empirical $P = 0.0009$), not on chromosome 16.

As evidence of linkage was already found in previous SLIC studies for the NWR and ELS traits, we considered whether the data from our independent sample met criteria for significant replication. A P value of 0.01 or lower is usually required to provide a confirmation of linkage at the 5% level (Lander & Kruglyak 1995). Using this threshold, linkage was confirmed on chromosome 16 for NWR and on chromosome 19 for ELS. While the NWR and ELS linkages only reached significance under one method of analysis (HE and DF, respectively), both models showed similar trends of results and the alternative analyses yielded empirical P values bordering on significance in both cases (NWR DF empirical $P = 0.015$, ELS HE empirical $P = 0.035$). For comparison, Fig. 4 displays the HE LOD scores for NWR and ELS obtained in the previous SLIC studies for the same region.

Discussion

There are a number of theories regarding the underlying basis of SLI (for a review, see Bishop 1997; Leonard 1998). Of

relevance here are two contrasting theories: limitations in phonological short-term memory capacity and delayed maturation of a specific linguistic brain system involved in the marking of grammatical morphology, in particular, finite verb inflections.

Gathercole & Baddeley (1989, 1990) argued that SLI may involve a specific deficit of phonological short-term memory. This component specializes in the temporary storage and processing of verbal material and, importantly, in their model, it is capacity limited. In SLI, it is proposed that this capacity is reduced, thus impeding efficient processing and storing of phonological information crucial to language learning. Gathercole & Baddeley (1990) found that children with SLI performed substantially below not only age controls but also chronologically younger language controls on a non-word repetition task (a task designed to measure phonological short-term memory), a finding supported by several subsequent studies (Conti-Ramsden & Durkin 2007; Dollaghan & Campbell 1998; Ellis Weismer *et al.* 2000; Montgomery 1995) and in languages other than English (Aguado *et al.* 2006; Reuterskiold-Wagner *et al.* 2005; Siu & Man 2006). That this appears to apply even in cases where the language problems have apparently been resolved (Conti-Ramsden *et al.* 2001) has provided a basis for suggesting that poor non-word repetition ability is not only a marker but also a key contributory trait of SLI.

A contrasting account, the 'Extended Optional Infinitive (EOI)' theory put forward by Rice (2000) and Rice *et al.* (1995), suggests that SLI results from slow maturation of the linguistic brain system involved in the grammatical marking of finiteness. While the grammatical marking system of a typically developing child matures relatively quickly, with substantial mastery by 5 years of age, children with SLI continue to treat finite marking as optional for an extended period of development.

Evidence for a genetic contribution to phonological short-term memory is mounting. Bishop *et al.* (1999) found that deficits in non-word repetition were highly heritable in a large

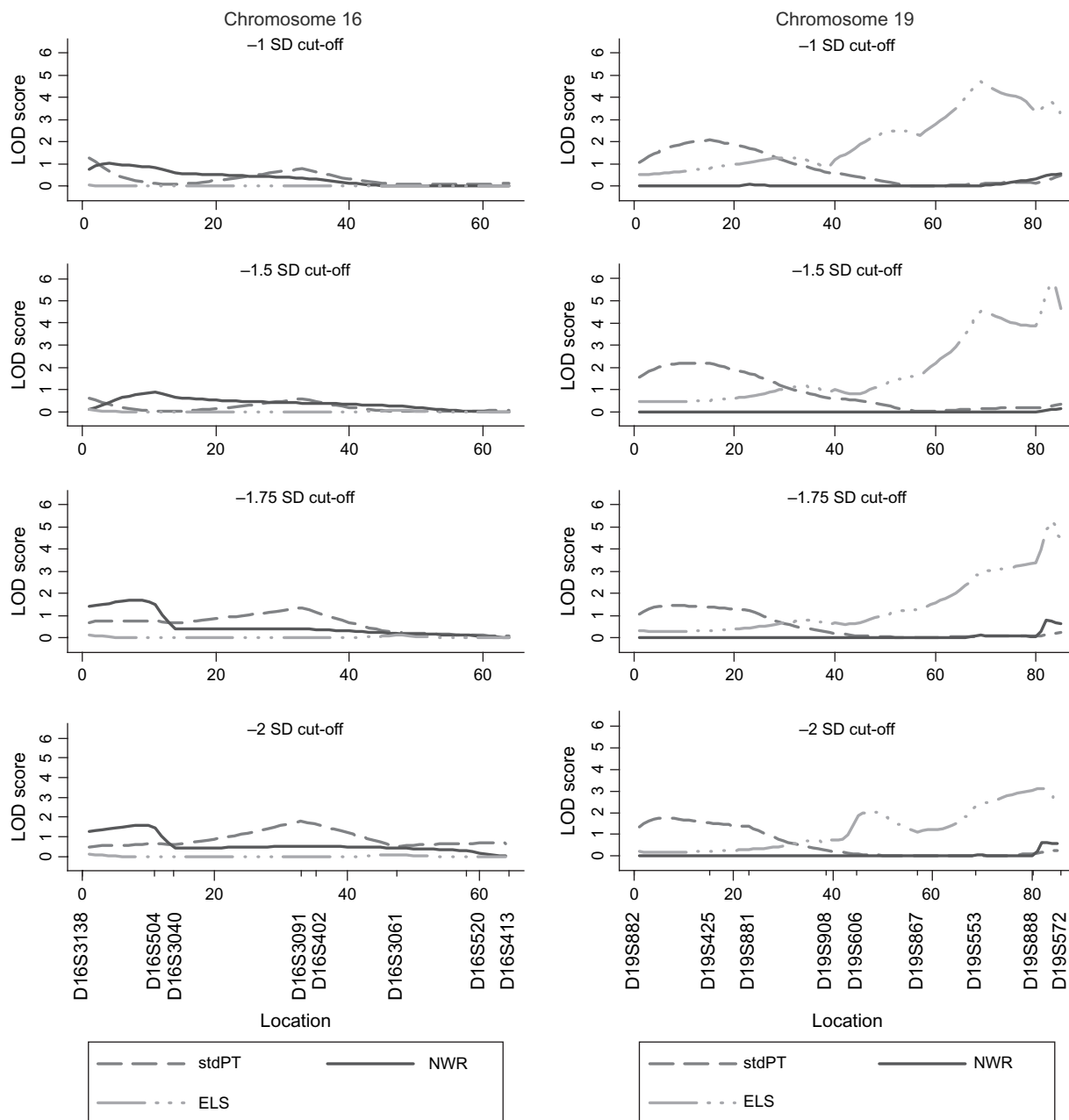


Figure 2: LOD scores along chromosomes 16 and 19 using DF models for several proband selection cut-offs (-1, -1.5, -1.75 and -2 SD).

sample of twins aged 7–12 years, and genome-wide linkage screens by The SLI Consortium (2002, 2004) have shown linkage to chromosomes 16 and 19 for non-word repetition among a range of language and literacy measures. Evidence for a genetic aetiology for the EOI theory of SLI is more slight. Rice *et al.* (1998) showed an excess of speech and language difficulties as well as language-related difficulties (e.g. reading) in first-degree relatives of children with SLI who had limitations in grammatical morphology. The twin analysis of

Bishop *et al.* (2006) found heritable components for both phonological short-term memory and grammatical morphology, but that these arose from largely distinct non-overlapping genetic effects.

In line with Bishop *et al.* (2006), our results suggest that phonological short-term memory is a good marker of heritable language impairment in SLI. We found familial aggregation of phonological short-term memory deficits as indexed by the non-word repetition task.

Table 2: Max LOD scores on chromosomes 16 and 19 obtained using DF models with different proband selection cut-offs

	Proband selection cut-off	Max LOD score (number of sibpairs)	
		Chromosome 16	Chromosome 19
std PT	-1 SD	1.26 (36)	2.08 (34)
	-1.5 SD	0.60 (34)	2.20 (32)
	-1.75 SD	1.34 (30)	1.46 (28)
	-2 SD	1.80 (28)	1.76 (26)
NWR	-1 SD	1.01 (24)	0.56 (23)
	-1.5 SD	0.90 (16)	0.14 (15)
	-1.75 SD	1.69 (13)	0.78 (12)
	-2 SD	1.58 (12)	0.63 (11)
ELS	-1 SD	0.05 (20)	4.72 (19)
	-1.5 SD	0.09 (17)	5.80 (16)
	-1.75 SD	0.12 (13)	5.26 (12)
	-2 SD	0.12 (12)	3.12 (11)

The number of sibpairs used in each DF analysis is reported within parentheses.

The picture for grammatical morphology as indexed by the PT task was more complex. Recall that our sample of probands was much older (mean age 14 years and 5 months) than those participating in previous studies (most below 7 years of age). Thus, age as a factor may have played an important role in the nature of our findings. Thus, unlike Bishop *et al.* (2006), we found no significant association between first-degree relatives' and probands' scores on the PT task when examined as a continuous variable. However, some evidence for familial aggregation was found for a binary measure of the PT task with a cut-off of -1.5 SD. Bishop (2003, 2005) has argued that structural aspects of language

such as grammatical morphology typically show little normal variation and have low ceiling levels, with children around 5 years of age showing considerable levels of mastery in these skills. We want to take this argument further and suggest that tense marking may not be a phenotypic trait that is measurable as a continuous dimension across development but instead may be a skill in which competence is either acquired or not acquired by early school age, comparable to a Piagetian stage in learning. In this sense, qualitative distinctions in the trait is what appears to be familial, while quantitative variability is likely to be more a consequence of non-familial factors, notably age and others, which may well be related to age, e.g. motivation/attention to task.

A recent account of SLI suggests that there may be a different interpretation of our findings. This account suggests that impairments in procedural memory may be implicated in the difficulties with grammatical morphology (Ullman 2001; Ullman & Pierpont 2005). Procedural memory is defined as the acquisition of new skills, both motor and cognitive, over multiple trials without the need of conscious awareness (Gabrieli 1998). Thus, the procedural memory system is well suited to learning probabilistic occurrences or relations, e.g. those required to access appropriate endings to inflect verbs (e.g. play-ed). It is of interest that research suggests that procedural memory may be an ability that is developmentally invariant (Rovee-Collier 1997; but see also Durkin 1994 and Thomas *et al.* 2004 for a more developmental perspective) and thus may be a skill where deficits are evidenced in a binary, all or none, fashion. Our findings on the PT task are consistent with both the linguistic interpretation provided by Rice *et al.* (1995) and the procedural memory deficit theory proposed by Ullman. However, taken together, our results suggest that deficits in multiple components at the same time might be necessary to produce SLI (see also Bishop 2006). In our sample, there were fewer cases of probands having

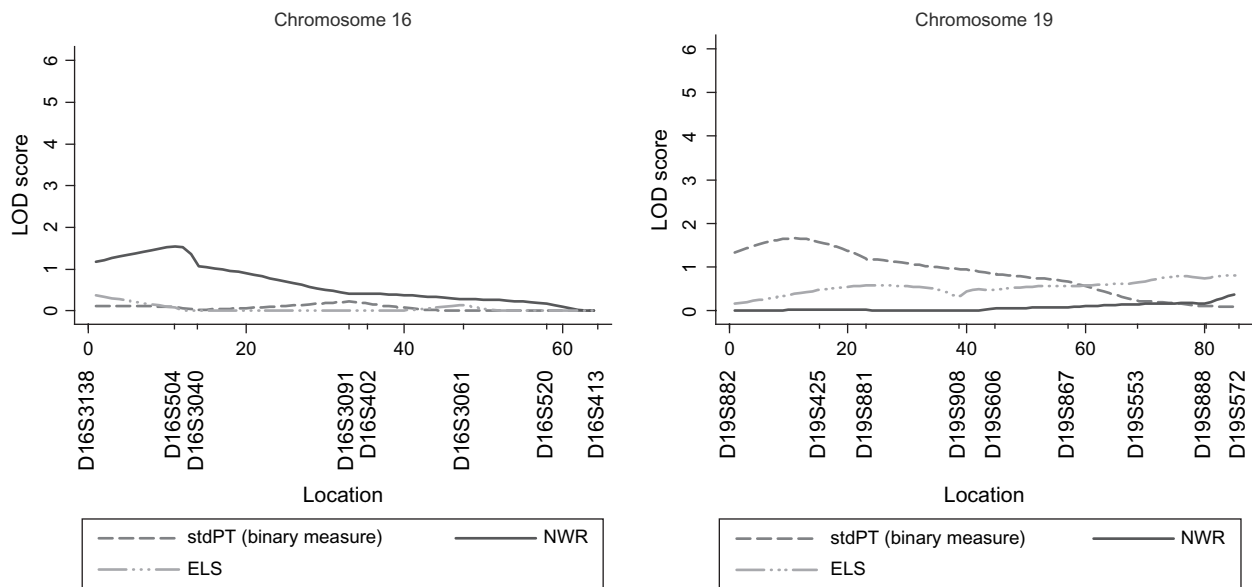


Figure 3: LOD scores obtained using HE models. For the PT binary measure, a logistic instead of a linear regression was fitted.

Table 3: Max LOD scores and their location and statistical significance along chromosomes 16 and 19 obtained using DF and HE models. One-sided empirical *P* values were computed using 100 000 permutations

	Chromosome 16			Chromosome 19		
	Max LOD	Location (cM from first marker)	Empirical <i>P</i> value*	Max LOD	Location (cM from first marker)	Empirical <i>P</i> value*
stdPT						
DF (continuous)	1.80	33	0.0128	2.20	12	0.0058
HE (binary)	0.23	33	0.1572	1.66	11	0.0009
NWR						
DF	1.69	8	0.0154	0.78	82	0.0339
HE	1.54	11	0.0023	0.36	85	0.0688
ELS						
DF	0.12	1	0.2340	5.80	83	0.0067
HE	0.37	1	0.0985	0.81	84	0.0352

*One-sided *P* values.

isolated deficits with PT or NWR; most cases presenting difficulties in both areas.

The results of our linkage analysis provide further support for previous SLIC findings on linkage to 16q for NWR and 19q

for ELS (The SLI Consortium 2002, 2004, Monaco & The SLI Consortium 2007). It should be borne in mind that it is not unusual to see variation between the exact positioning of linkage peaks between samples. On chromosome 16, the

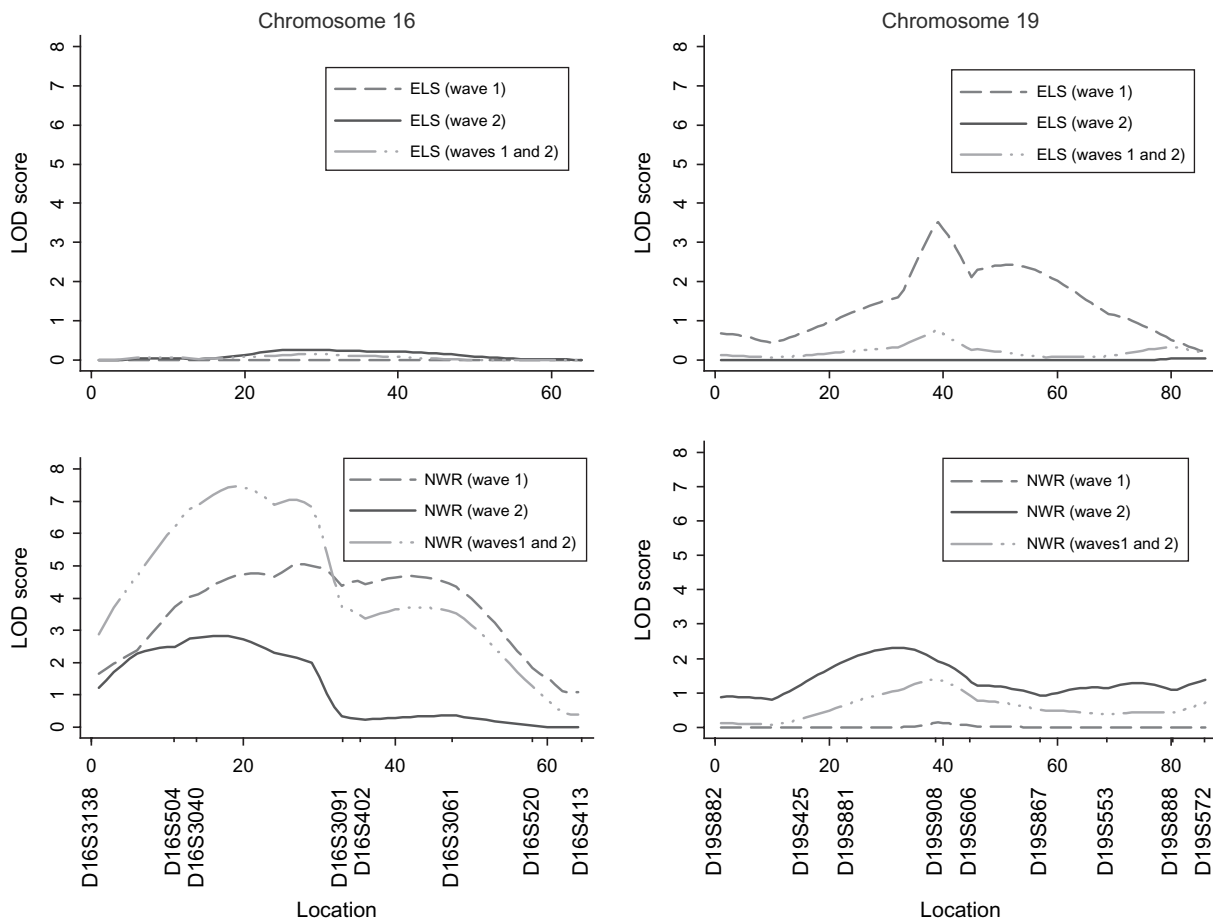


Figure 4: Haseman–Elston LOD scores for NWR and ELS as obtained by the SLIC for the same regions of our plots.

peak of linkage for NWR in this new sample was found in a position close to the region for which SLIC previously reported evidence of linkage. For ELS on chromosome 19, while the new sample also replicates the presence of linkage, the peak does not coincide with that found by SLIC at about 35–40 cM on the map used in this paper. However, the previous SLIC peak was made up from the combined analysis of several samples and the peaks from individual samples were distributed across the entire region. Furthermore, while the 40 cM region in this study is not the peak of linkage, linkage in that region is not completely flat and exceeds a LOD score of 1. Hence, while these analyses do not preclude the existence of two influential genes within this region, set in this context of sample heterogeneity, we do not consider the evidence to be strong. As such, we suggest that this should be treated as a single region of linkage until specific genetic variants are identified at which time the possibility of a second genetic variant can be explored.

In addition to replication findings, we report new findings for the PT measure, for which to our knowledge there has been no previous molecular genetic study. When PT was measured on the continuous scale, the DF models suggested linkage on both chromosomes but evidence for linkage was stronger on chromosome 19. When a binary measure for PT (cut-off of -1.5 SD) was considered, linkage was only observed on chromosome 19, not on chromosome 16. This is consistent with there being distinctive genetic bases to PT and NWR and similar to the Bishop *et al.* (2006) twin analysis finding of limited sharing of additive genetic variance for grammar and phonological short-term memory. However, in the context of the heterogeneity of findings that we often see across samples and measures, we again do not consider this strong evidence. Nonetheless, these results suggest that the *SLI2* region on chromosome 19 is worthy of further investigation using larger samples and a more extended and refined range of measures.

References

- Adams, A.-M. & Gathercole, S.E. (1995) Phonological working memory and speech production in pre-school children. *J Speech Hear Res* **38**, 403–414.
- Adams, A.-M. & Gathercole, S.E. (1996) Phonological working memory and spoken language development in young children. *Q J Exp Psychol* **49A**, 216–233.
- Aguado, G., Cuetos-Vega, F., Domezain, M. & Pascual, B. (2006) Repetition of pseudo-words in Spanish children with specific language disorder: a psycholinguistic marker. *Rev Neurol* **43** (Suppl. 1), S201–S208.
- Amos, C.I., Elston, R.C., Wilson, A.F. & Bailey-Wilson, J.E. (1989) A more powerful robust sib-pair test of linkage for quantitative traits. *Genet Epidemiol* **6**, 435–449.
- Bishop, D.V.M. (1997) *Uncommon Understanding: Development and Disorders of Language Comprehension in Children*. Psychology Press, Hove.
- Bishop, D.V.M. (2003) Genes and language. In Krude, T. (ed.), *DNA: Changing Science and Society*. Cambridge University Press, Cambridge, pp. 144–165.
- Bishop, D.V.M. (2005) DeFries-Fulker analysis of twin data with skewed distributions: cautions and recommendations from a study of children's use of verb inflections. *Behav Genet* **35**, 479–490.
- Bishop, D.V.M. (2006) What causes specific language impairment in children. *Curr Dir Psychol Sci* **15**, 217–221.
- Bishop, D.V.M., Bishop, S.J., Bright, P., James, C., Delaney, T. & Tallal, P. (1999) Different origin of auditory and phonological processing in children with language impairment: evidence from a twin study. *J Speech Lang Hear Res* **42**, 155–168.
- Bishop, D.V.M., Adams, C.V. & Norbury, C.F. (2006) Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes Brain Behav* **5**, 158–169.
- Blackwelder, W.C. & Elston, R.C. (1982) Power and robustness of sib-pair linkage test and extension to larger sibships. *Commun Stat Theor Meth* **11**, 449–484.
- Carey, G. & Williamson, J. (1991) Linkage analysis of quantitative traits: increased power by using selected samples. *Am J Hum Genet* **49**, 786–796.
- Conti-Ramsden, G. & Botting, N. (1999a) Characteristics of children attending language units in England: a national study of 7-year old. *Int J Lang Commun Disord* **34**, 359–366.
- Conti-Ramsden, G. & Botting, N. (1999b) Classification of children with specific language impairment: longitudinal considerations. *J Speech Lang Hear Res* **42**, 1195–1204.
- Conti-Ramsden, G. & Durkin, K. (2007) Phonological short-term memory, language and literacy: developmental relationships in early adolescence in young people with SLI. *J Child Psychol Psychiatry* **48**, 147–156.
- Conti-Ramsden, G., Crutchley, A. & Botting, N. (1997) The extent to which psychometric tests differentiate subgroups of children with specific language impairment. *J Speech Lang Hear Res* **40**, 765–777.
- Conti-Ramsden, G., Botting, N. & Faragher, B. (2001) Psycholinguistic markers for specific language impairment (SLI). *J Child Psychol Psychiatry* **42**, 741–748.
- Dollaghan, C. & Campbell, T. (1998) Nonword repetition and child language impairment. *J Speech Lang Hear Res* **41**, 1136–1146.
- Durkin, K. (1994) The lure and language of implicit memory: a developmental perspective. In Ellis, N. (ed.), *Implicit and Explicit Learning of Languages*. Academic Press, London, UK, pp. 523–548.
- Ellis Weismer, S., Tomblin, J.B., Xuyang, Z., Buckwalter, P., Chynoweth, J.G. & Jones, M. (2000) Nonword repetition performance in school-age children with and without language impairment. *J Speech Lang Hear Res* **43**, 865–878.
- Elston, R.C. (1998) Methods of linkage analysis and the assumptions underlying them. *Am J Hum Genet* **63**, 931–934.
- Elston, R.C. & Cordell, H.J. (2001) Overview of model-free methods for linkage analysis. *Adv Genet* **42**, 135–150.
- Fulker, D.W., Cardon, L.R., DeFries, J.C., Kimberling, W.J., Pennington, B.F. & Smith, S.D. (1991) Multiple regression analysis of sib-pair data on reading to detect quantitative trait loci. *Reading Writing Interdiscip J* **3**, 299–313.
- Gabrieli, J.D.E. (1998) Cognitive neuroscience of human memory. *Annu Rev Psychol* **49**, 87–115.
- Gathercole, S.E. & Baddeley, A.D. (1989) Evaluation of the role of STM in the development of vocabulary in children: a longitudinal study. *J Mem Lang* **29**, 336–360.
- Gathercole, S.E. & Baddeley, A.D. (1990) Phonological memory deficits in language disordered children: is there a causal connection? *J Mem Lang* **29**, 336–360.
- Gathercole, S.E. & Baddeley, A.D. (1996) *Nonword Memory Test*. Centre for Working Memory and Learning, The University of York, UK. URL: <http://www.york.ac.uk/res/wml/test/wm.htm>.
- Gayán, J., Smith, S.D., Cherny, S.S., Cardon, L.R., Fulker, D.W., Brower, A.M., Olson, R.K., Pennington, B.F. & DeFries, J.C. (1999) Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *Am J Hum Genet* **64**, 157–164.
- Grigorenko, E.L., Wood, F.B., Meyer, M.S. & Pauls, D.L. (2000) Chromosome 6p influences on different dyslexia-related cognitive processes: further confirmation. *Am J Hum Genet* **66**, 715–723.
- Haseman, J.K. & Elston, R.C. (1972) The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* **2**, 3–19.
- Huber, P.J. (1967) The behaviour of maximum likelihood estimators under non-standard conditions. In *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, CA, pp. 221–233.

- Knopik, V.S., Smith, S.D., Cardon, L., Pennington, B., Gayán, J., Olson, R.K. & DeFries, J.C. (2002) Differential genetic etiology of reading component processes as a function of IQ. *Behav Genet* **32**, 181–198.
- Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S.A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S.T., Frigge, M.L., Thorgeirsson, T.E., Gulcher, J.R. & Stefansson, K. (2002) A high-resolution recombination map of the human genome. *Nat Genet* **31**, 241–247.
- Kruglyak, L., Daly, M.J., Reeve-Daly, M.P. & Lander, E.S. (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* **58**, 1347–1363.
- Lander, E. & Kruglyak, L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* **11**, 241–247.
- Leonard, L.B. (1998) *Children with Specific Language Impairment*. MIT Press, London, UK.
- Marchman, V.A., Wulfeck, B. & Ellis Weismer, S. (1999) Morphological productivity in children with normal language and SLI: a study of the English past tense. *J Speech Lang Hear Res* **42**, 206–219.
- Monaco, A.P. & The SLI Consortium (2007) Multivariate linkage analysis of Specific Language Impairment (SLI). *Ann Hum Genet* **71**, 660–673.
- Montgomery, J.W. (1995) Sentence comprehension in children with specific language impairment: the role of phonological working memory. *J Speech Hear Res* **38**, 187–199.
- Neils, J. & Aram, D.M. (1986) Family history of children with developmental language disorders. *Percept Mot Skills* **63**, 655–658.
- Rabe-Hesketh, S., Skrondal, A. & Pickles, A. (2002) Reliable estimation of generalized linear mixed models using adaptive quadrature. *Stata J* **2**, 1–21.
- Reuterskiold-Wagner, C., Sahlen, B. & Nyman, A. (2005) Non-word repetition and non-word discrimination in Swedish preschool children. *Clin Linguist Phon* **19**, 681–699.
- Rice, M.L. (2000) Grammatical symptoms of specific language impairment. In Bishop, D.V.M. & Leonard, L.B. (eds), *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome*. Psychology Press, Hove, pp. 17–34.
- Rice, M.L., Wexler, K. & Cleave, P.L. (1995) Specific language impairment as a period of extended optional infinitive. *J Speech Hear Res* **38**, 850–863.
- Rice, M.L., Haney, K.R. & Wexler, K. (1998) Family histories of children with SLI who show extended optional infinitives. *J Speech Lang Hear Res* **41**, 419–432.
- Rovee-Collier, C. (1997) Dissociations in infant memory: rethinking the development of implicit and explicit memory. *Psychol Rev* **104**, 467–498.
- Semel, E., Wiig, E. & Secord, W. (1987) *Clinical Evaluation of Language Fundamentals – Revised*. The Psychological Corporation, San Antonio, TX.
- Shete, S., Jacobs, K.B. & Elston, R.C. (2003) Adding further power to the Haseman and Elston method for detecting linkage in larger sibships: weighting sums and differences. *Hum Hered* **55**, 79–85.
- Simkin, Z. & Conti-Ramsden, G. (2001) Non-word repetition and grammatical morphology: normative data for children in their final year of primary school. *Int J Lang Commun Disord* **36**, 395–404.
- Siu, E. & Man, D.W. (2006) Working memory and sentence comprehension of Hong Kong Chinese children with specific language impairment. *Int J Rehabil Res* **29**, 267–269.
- The SLI Consortium (2002) A genomewide scan identifies two novel loci involved in specific language impairment. *Am J Hum Genet* **70**, 384–398.
- The SLI Consortium (2004) Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *Am J Hum Genet* **74**, 1225–1238.
- StataCorp (2005) *Stata Statistical Software: Release 9.0*. Stata Corporation, College Station, TX.
- Tallal, P., Ross, R. & Curtiss, S. (1989) Familial aggregation in specific language impairment. *J Speech Hear Disord* **54**, 167–173.
- Thomas, K.M., Hunt, R.H., Vizueta, N., Sommer, T., Durston, S., Yang, Y. & Worden, M.S. (2004) Evidence of developmental differences in implicit sequence learning. *J Cogn Neurosci* **16**, 1339–1351.
- Tobin, J. (1958) Estimation of relationships for limited dependent variables. *Econometrica* **26**, 24–36.
- Ullman, M.T. (2001) The declarative/procedural model of lexicon and grammar. *J Psycholinguist Res* **30**, 37–69.
- Ullman, M.T. & Pierpont, E.I. (2005) Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex* **41**, 399–433.

Acknowledgments

The authors gratefully acknowledge the support of the Wellcome Trust (Grant 060774) to G.C.-R. and Principal Research Fellowship to A.P.M. S.E.F. is a Royal Society Research Fellow. Thanks go to Helen Betteridge, Emma Knox and Catherine Pratt for their help with data collection and James Cleak for the DNA preparation. The authors would also like to thank the families who helped us with this research.