Published in final edited form as:

J Child Psychol Psychiatry. 2009 January ; 50(0): 116–125. doi:10.1111/j.1469-7610.2008.02006.x.

Speaking Genes or Genes for Speaking? Deciphering the Genetics of Speech and Language

Elena L. Grigorenko

Yale University

Abstract

Background—This article selectively reviews the status of the genetic research in the field of speech and language disorders.

Methods—Major contributions to the field are selected, presented, and discussed.

Results—The field presents itself through a variety of findings, characterized by both consistencies and inconsistencies.

Conclusions—The last 30+ years of the field unequivocally testify to the importance of the genetic factors in the acquisition of speech and language. However, the details of how these factors exert their influence are yet to be determined.

As of June 22, 2008, we were 4,156 short of an estimated 1,000,000 words in the English language (<http://www.languagemonitor.com/>). Against the backdrop of new words emerging and old words fading away, every person for whom English is a native language or who tries to acquire English as a foreign tongue is faced with the monumental task of mastering a bulk of them. Regardless of whether native or foreign, everyone agrees that people are not born with explicit knowledge of any languages' vocabularies (also referred as lexicons) and the rules by which words are woven together to make coherent speech. A newborn child does not know, pronounce, or manipulate any words. However, a child entering school has a lexicon of ~13,000 words which, by the end of schooling, expands to ~60,000 words and during adulthood reaches ~120,000 words (Pinker, 1999).

How this ascent from 0 to 120,000 (and for a tiny minority, all the way to close to a million!) occurs and how people acquire the rules of making meaningful sounds, whirling them into words, and assembling endless combinations of words into meaningful sentences have been central questions in the fields of developmental and cognitive psychology, linguistics, and speech and language pathology since their emergence. This field of study within and across disciplines is typically referred to as language acquisition. And although everyone agrees that language acquisition has distinct developmental stages, there is also consensus that a huge amount of individual differences characterize this process. These individual differences appear to be quantitatively distributed for all facets of language acquisition (i.e., from mastering sounds to comprehending complex oral and written inputs), with individuals scoring at the lower tail of the distribution for one or more facets of language acquisition, demonstrating substantial difficulties in mastering language that are so

significantly different from trajectories of typical development that they exemplify atypical development or different pathologies of speech and language. Mounting evidence indicates that a very important source of these individual differences, across the whole distribution but particularly at its lower tail, is the genome. The goal of this brief essay is to comment on behavior- and molecular-genetic studies that elucidate the role of the genome in the atypical acquisition of language that is manifested by disorders of speech and language.

Phenotypes: Typologies and Facets

In fact, children with atypical patterns of language acquisition form a sizeable group. Specific difficulties characteristic of such patterns have been observed and studied in children for a long time, but presently, there is no clear classification of the related difficulties (Verhoeven & van Balkom, 2004). In fact, there are many alternative classifications, some of which are institutionalized in diagnostic manuals or conventionally used in research and clinical practice. For example, below are categories used in the world's two leading classification manuals, published by the World Health Organization (ICD-10, <http://www.who.int/classifications/icd/en/>) and the American Psychiatric Association (DSM-IV, <http://www.psychiatryonline.com/resourceTOC.aspx?resourceID=1>), respectively. Specifically, the following categories are listed in the corresponding manuals.

- ICD-10: ICD-10's Chapter V (*Mental and Behavioural Disorders*)
 - Disorders of Psychological Development
 - *Specific Developmental Disorder of Speech and Language*
(Specific Speech Articulation Disorder, Expressive Language Disorder, Receptive Language Disorder, Acquired Aphasia with Epilepsy, Other Developmental Disorders of Speech and Language, and Developmental Disorder of Speech and Language, Unspecified)
- DSM-IV: *Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence*
 - Communication Disorders* (Expressive Language Disorder, Mixed Receptive–Expressive Language Disorder, Phonological Disorder, Stuttering, and Communication Disorder NOS)

Even a cursory examination of these lists suggests substantial uncertainty and disagreement in the field. Moreover, as mentioned earlier, the research literature does not directly map itself onto either diagnostic manual. Here, I capitalize on the distinctions made in the research literature between disorders of speech and sound (Speech and Sound Disorders, SSD) and disorders of language (Specific Language Impairment, SLI). In presenting behavior- and molecular-genetic findings, I use these two categories broadly, including in SSD such disorders as apraxia and stuttering, and in the SLI category both receptive and expressive language disorders. Finally, I will collectively refer to all these disorders as Disorders of Speech and Language (DSL).

Three additional remarks are in order before I review corresponding behavior- and molecular-genetic findings on DSL.

First, psychological and linguistic theories of language acquisition assume that, when unfolding typically, this process assumes the formation of a number of types of representations (see Text Box 1). These representations form “facets” of language acquisition, and developmental deficits might arise in one, some, or all of them, depending on the type and severity of DSL. A fundamental question in studies of language acquisition is how all these psychological representations are acquired and what goes wrong in the formation of these representations that leads to the development of a DSL. Various DSL are characterized by patterns of deficits in phonological, morphological, lexical, semantic, syntactic, and pragmatic facets, so that there are DSL-general (e.g., in phonological representation) and DSL-specific (e.g., in practice) deficits. Correspondingly, different phenotypes can be derived from assessments that tap into these facets of language acquisition.

Second, as is the case for many other common complex disorders, there is an ongoing debate about how to define DSL: through their categorical distinctness as defined by a bimodal distribution of some underlying liability trait, or through their representation of the extreme lower end of normal developmental variation where the distribution does not separate typical and atypical language development (Bishop, 1994b, 2005; Cole, Schwartz, Notari, Dale, & Mills, 1995). This debate is central to assumptions about the specificity versus generality of the genetic mechanisms affecting typical and atypical (disordered) language acquisition. Specifically, if the distributions of typical and atypical skills are assumed to be discrete, then different genetic mechanisms¹ might be assumed to contribute to the acquisition of different skills. If disordered skills are viewed as the lower end of the normal distribution of language acquisition skills, then the same genetic mechanisms can be assumed to influence the manifestation of both typical and atypical skills.

Third, although researchers may disagree on the characterization of the underlying linguistic–cognitive problems, their findings on the manifestations of the disorders are quite consistent. Thus, children with SSD demonstrate inadequate mastery of sound production and perception (Shriberg, 2002). Children with SLI demonstrate great difficulty acquiring and using grammatical markers that express purely structural relations (the so-called functional categories), such as various tense and agreement markers, including the past tense *–ed*, the third person singular *–s*, the auxiliary verb *do*, and so on. Yet, there are multiple theories attempting to explain specific disorders. For example, intending to account for SLI, researchers have suggested that this impairment is (1) a specific distortion of syntax acquisition (Rice, Wexler, & Cleave, 1995); (2) a result of poor auditory perception (Bates,

¹In fact, hypothetically, this “line of reasoning” can be also applied to different mutations within a single gene, with these mutations resulting in different outcomes. For example, a null mutation (i.e., a mutation that leads to the non-synthesis of the protein coded by a mutated gene or to the synthesis of a defective protein lacking its typical functional properties) in a gene contributing to genetic variation in language acquisition might result in the formation of a dichotomous distribution with individuals who possess such a mutation not able to acquire language at all. A less deleterious mutation (i.e., one that challenges some but preserves other functions of a protein) might be associated with a less severe phenotype. Correspondingly, if many mutations of different degrees of deleteriousness exist within one gene, these various mutations (when, of course, considered, in a very large sample of individuals) might result in a more continuous distribution.

2004); (3) a consequence of phonological deficits (Joanisse, 2004) or (4) a deficit in processing capacity (Bishop, 1994a); or some combination of all these factors (Leonard, 1998). Of interest to this discussion is that behavior-genetic research has been used to clarify the theoretical and phenomenological work surrounding DSL. For example, in one study, researchers (Bishop, Adams, & Norbury, 2006) attempted a direct comparison of selected theories by measuring indices of phonological processing (short-term phonological memory) and grammaticality (a combined index of verb inflections) in a set of twins, ~66% of whom were selected because of their language difficulties and ~33% of whom were chosen at random from a large sample with no evidence of a language disorder. Both measures provided high estimates of heritability for poor performance (.61 and .74, for phonological and grammatical indicators, respectively). Bivariate analyses revealed no indication of genetic overlap between the two measures, suggesting that they were probably driven by different genetic etiologies. This finding is challenging to explain within the context of theories that view SLI as an outcome of a phonological deficiency of some kind.

In summary, given the debates in the field regarding general mechanisms of language acquisition and specifics on the definition, manifestation, typologies, and theories of different DSL, the task of defining phenotypes in behavior- and molecular-genetic studies of DSL is critical. Therefore, in the discussion below I distinguish between specific phenotypes when these distinctions are important for the discussion.

An Overview of Behavior- and Molecular-Genetic Studies of DSL

Familiality and heritability of DSL

There is substantial evidence that DSL cluster in families. Investigations of this clustering have been carried out using a number of methodologies, including biological [e.g., (Choudhury & Benasich, 2003)], adoptive family [e.g., (Felsenfeld & Plomin, 1997)], and twin designs [e.g., (Oliver & Plomin, 2007)]. Evidence from these studies has clearly demonstrated the importance of genetic factors in the development of various forms of DSL (e.g., stuttering, SSD, SLI).

Specifically, Stromswold (1998) reviewed a number of studies of familial aggregation of DSL. These studies indicated a higher than expected presence of DSL among the relatives of DSL probands: The median incidence rate for language difficulties in the families of children with language impairment was 35%, compared with a median incidence rate of 11% in control families (Stromswold, 1998). The reported prevalence estimates range from 24 to 78%, ~5 to 10 times higher than the 2 to 7% observed in the general population.

Twin studies have consistently reported a substantial difference in concordance between monozygotic and dizygotic twins, with the former showing higher concordance rates than the latter [ranging from 25 to 100% in published literature (Stromswold, 2001)], thus, indicating the presence of genetic influences. In addition, heritability estimates are high for a variety of psychometric indicators capturing psychological processes or facets thought to be deficient in DSL (DeThorne et al., 2006; Kovas et al., 2005). Of note is that heritability estimates appear to be high whether ascertained directly from the child through

psychometric assessment or clinical interview or through parent or teacher report (Bishop, Laws, Adams, & Norbury, 2006).

Considered holistically, the accumulated evidence points to the substantial role of genes and the relatively small role of environment in the development and manifestation of DSL. Yet, collectively these studies have not revealed particular replicable modes of genetic transmission for any of the DSL, although multiple attempts have been made. Presently, it is hypothesized that, as is the case for many common complex disorders, the etiology of DSL is probably related to linear and interactive influences of combinations of gene and environmental factors (Smith, 2007).

The KE family

A very special and interesting case of a clear pattern of the transmission of a DSL is document in a large, three-generation pedigree from the United Kingdom (referred to as KE), in which about 50% of the members suffer from a severe complex disorder that distorts speech, language, and cognitive functioning and is passed through the generations in a simple Mendelian pattern of dominant transmission (i.e., the disorder is present in each generation at a particular rate). Since the family was first presented in the literature (Hurst, Baraitser, Auger, Graham, & Norell, 1990), many subsequent steps have been undertaken in tracking the manifestation [e.g., (Watkins, Dronkers, & Vargha-Khadem, 2002)], brain correlates [e.g., (Belton, Salmond, Watkins, Vargha-Khadem, & Gadian, 2003)], and genetic etiology of the disorder [e.g., (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001)].

As mentioned earlier, this family's phenotype is complex, with deficiencies manifesting as verbal dyspraxia (i.e., speech articulation problems) and difficulties in receptive and expressive vocabulary, morphology, and syntax. Also of note is that, in the KE family, difficulties penetrate all types of linguistic processing (i.e., oral and written language) in aspects of both reception and production. Finally, when compared with the unaffected individuals in the pedigree as a group, affected individuals are characterized by lower nonverbal IQ. However, the cognitive difficulties do not co-segregate “true” with speech and language difficulties, and this might suggest that two independent disorders segregate in this family. Thus, the phenotypic texture of difficulties transmitted in this pedigree is complex. Moreover, a number of deficits (e.g., cognitive deficits) are observed among unaffected individuals. Yet, there appears to be a single phenotype that best differentiates affected from unaffected individuals—individual performance on a nonword repetition task, which is a measure of phonological representation (Watkins et al., 2002).

Given the dominant pattern of transmission of this complex disorder in the KE family, researchers expected that the disorder was controlled by a single major gene [for a more complete review of the genetic etiology of the disorder in the KE family, see (Fisher, Lai, & Monaco, 2003)]. Through multiple steps, researchers identified a probable single gene, located on chromosome 7q31. Bioinformatics yielded the discovery that the gene encoded a DNA-binding motif, a three-dimensional structural element or fold within the DNA sequence, which also appeared in a variety of other molecules, in the so-called forkhead/ winged-helix domain (Kaufmann & Knöchel, 1996). Correspondingly, the gene was named *FOXP2* (forkhead box P2). This family of genes codes for proteins that specifically bind to

promoter or enhancer regions on DNA and thereby control the transcription of other genes. When the gene was sequenced in a member of the KE family, a mutation divulging a G-to-A (histidine to arginine) nucleotide change in exon 14 of *FOXP2* was discovered. This mutation was not found in any individual in a group of 364 screened control subjects without DSL, but co-segregated perfectly with the affected status in the KE family and, thus, was recognized as a cause of the disorder transmitted in this family (Lai et al., 2001). Subsequent investigations of other patients with similar phenotypes revealed additional mutations in *FOXP2* that caused deficiencies in these patients (MacDermot et al., 2005). It was hypothesized that, like other transcription factors, the FOXP2 protein might be important in embryogenesis, especially at the stage of brain development (Fisher et al., 2003). However, the specifics of this causal chain have yet to be determined (French et al., 2007; White, Fisher, Geschwind, Scharff, & Holy, 2006).

Molecular-genetic studies of DSL

The identification of *FOXP2* as a causal gene for DSL in the KE family and four additional cases of complex DSL generated significant and enthusiastic interest in this gene. A number of researchers began to investigate genetic associations between *FOXP2* and different types of DSL. These investigations unfolded to include more common cases of speech and language impairments (Meaburn, Dale, Craig, & Plomin, 2002; Newbury et al., 2002; O'Brien, Zhang, Nishimura, Tomblin, & Murray, 2003) and autism (Bartlett et al., 2002; Newbury et al., 2002). At present, there is no consistent evidence supporting the involvement of *FOXP2* in forms of DSL more common than those in the KE family and a few other interesting cases.

Conscious of the apparent noninvolvement of *FOXP2* in DSL in the general population, researchers carried out a number of molecular-genetic studies of DSL of two different types: (1) whole-genome screens and (2) regional studies with various forms of DSL. Using whole-genome scans, researchers attempted to identify regions of the genome that might contain susceptibility genes for DSL anew, whereas in the regional studies the task was to test for linkage with DSL those regions that had been implicated in comorbid developmental disorders such as autism and developmental dyslexia. Here both types of studies are illustrated briefly.

Genome-wide screens of DSL phenotypes—Seven genome-wide scans will be commented on here (see Table 1).

Five large extended Canadian families of Celtic ancestry, initially ascertained for a study of schizophrenia but in which multiple individuals met criteria of SLI, were genotyped with a set of highly polymorphic markers spaced at ~9cM apart (Bartlett et al., 2002). Three categorical phenotypes were used in these analyses: (1) language impairment (spoken language quotient standard score of 85); (2) reading impairment (1 SD discrepancy between performance IQ and nonword reading ability); and (3) clinical impairment (poor performance on language or reading subtests and/or at least two years of therapeutic intervention). The outcomes of this analysis identified a number of regions of interest, but the results varied by phenotype and genetic model.

Another whole-genome scan (density at $\sim 8\text{cM}$) was carried out with 219 sibpairs from 98 small UK nuclear families ascertained through a proband with SLI (The SLI Consortium, 2002). In this study, the phenotypes were captured by quantitative traits, nonword repetition and receptive and expressive scales of a standardized clinical assessment battery of language functioning. The results revealed two susceptibility regions, 16q24 (nonword repetition) and 19q13 (expressive language), in order of significance. Each of the identified regions spans $\sim 30\text{cM}$ and potentially contains ~ 100 genes (Haines & Camarata, 2004). To follow up on previously reported findings, the researchers ascertained an additional sample of 86 new families; both regions were linked in this sample as well [however, both regions 16q and 19q appeared to be linked to nonword repetition (The SLI Consortium, 2004)]. Furthermore, when both samples were amalgamated (393 sib pairs from 184 families), the 16q peak reached the threshold referred to as highly significant ($MLS = 7.46$ and 4.41 for multi- and single-point, respectively). However, the 19q results were not replicated in the amalgamated sample, which might not be that surprising given both samples' link to two different phenotypes. It is also curious that, although smaller in magnitude, linkage in the same region on 16q was established with three reading-related measures (basic reading, spelling, and reading comprehension). A recent multivariate re-analysis of the data from the combined sample, in which the same set of genetic data but an extended set of the phenotypic data were used, supported the involvement of both 16q and 19q regions and introduced a new region on 10q (Monaco & The SLI Consortium (SLIC), 2007).

Five whole-genome scans have been carried out to detect regions harboring genes controlling genetic variation in stuttering, a developmental disorder of speech characterized by disfluency of verbal expression manifested in involuntary repetitions of sounds or syllables (see Table 1 for detail). Three whole-genome scans were completed at the density of $\sim 10\text{cM}$ with highly polymorphic markers (STRPs). The fourth whole-genome scan used a panel of 10,000 diallelic polymorphisms (SNPs) genotyped in 100 families from The Illinois International Genetics of Stuttering Project, and revealed general linkage peaks on chromosomes 9q and 15q for the whole sample, sex-specific signals on chromosomes 7q for male- and 21q for female-only data subsets, as well as evidence for gene $\times$ gene interactions involving chromosomes 12q and 7q and 2q and 7q and 9q (Suresh et al., 2006). Finally, a genome screen was conducted on a sample of individuals with stutters connected in a single 232-person Hutterite family: $\sim 1,200$ markers, both STRPs and SNPs, were used (Wittke-Thompson et al., 2007). Three linkage signals were detected (3q, 13q, and 15q), though none of them were particularly strong. All five investigations used only categorical diagnoses-based phenotypes, although more than one (e.g., narrow and broad, susceptibility and persistence) diagnostic category was used.

In summary, although each particular study produced an array of results, there were no obvious overlaps of chromosomal regions in the primary linkage analyses. Such a pattern of results might indicate that (1) many different genes are involved in the development and manifestation of DSL; (2) DSL are very heterogeneous disorders and each sample or a disorder might have its own genetic etiology; (3) no study, when considered separately, has enough power to produce consistent results and some of these signals might still be false positives; (4) a higher density of genetic markers (e.g., as in the study by Suresh et al., 2006 and higher) is needed to detect "weak" signals consistently; and (5) the phenotypic

definitions of deficits segregating in families of the probands of DSL might not be precise enough. All these possible explanations of the inconsistency of the current pattern of results might be valid; therefore, further molecular-genetic studies of DSL are needed to explain and understand this pattern.

Regional genomic studies of DSL phenotypes—In addition to genome-wide screens, a number of region-specific studies with a variety of DSL phenotypes have been carried out. For example, in a large sample of families of probands with SLI (O'Brien et al., 2003), the 7q31 region was covered with 6 markers in and around the *FOXP2* gene. Although no genetic association with *FOXP2* was established, strong association was found with a marker within the *CFTR* (cystic fibrosis transmembrane conductance regulator) gene and the marker D7S3052, both adjacent to *FOXP2*.

A different research group investigated a set of susceptibility regions identified in studies of autism. Capitalizing on genetic findings in the field of autism and on a set of results from their own whole-genome screen (Bartlett et al., 2002), the researchers studied two SLI samples (one from Canada and one from the US) for linkage to the previously identified regions on chromosomes 2, 7, and 13 (Bartlett et al., 2004). The formerly registered linkage to 13q21 was replicated in both samples. In addition, while the overall evidence for linkage signals on chromosomes 2 and 7 was weak, markers in the *CFTR* gene (see above) demonstrated genetic association with the diagnosis of SLI.

Following up on a lead obtained through studies of developmental dyslexia, a different group of researchers explored linkage to the pericentrometric region of chromosome 3 in 77 families ascertained through a child with SSD (Stein et al., 2004). In this study, phenotypes were defined as quantitative traits capturing a number of linguistic processes (i.e., phonological memory, phonological representation, articulation, receptive and expressive vocabulary, and reading decoding and comprehension skills). Overall, the results indicated the linkage of a number of the investigated phenotypes to the region on chromosome 3, with the highest linkage score found for measures of phonological memory, followed by scores for single-word decoding for both real and pseudowords (nonwords). Using the same but enlarged sample (151 families), the same group of investigators also studied the 15q14–q21 region that had been previously implicated in a number of studies of developmental disorders including autism and developmental dyslexia. A categorical phenotype of SSD and a number of SSD-related continuous phenotypes were used in these analyses. Positive linkage signals were obtained in the 15q14 region for the phenotypes of SSD, and quantitative phenotypes of oral–motor function, articulation, and phonological memory (Stein et al., 2006). Yet another region, previously implicated as a candidate region for developmental dyslexia, 1p34–36, was investigated in the same sample with two categorical diagnoses (SSD and SLI) and a number of DSL-related phenotypes. Globally speaking, the linkage to 1p34–36 was replicated, though the specific peaks varied for different phenotypes. The strongest signals were obtained for articulation and listening comprehension, although the precise locations of these signals did not overlap (Miscimarra et al., 2007).

Still another set of region-specific studies was carried out on another sample of 86 probands with SSD and their siblings from 65 families (Smith, Pennington, Boada, & Shriberg, 2005). Seven phenotypes were used—five indicators of speech development and two spoken language measures of phonological processing (nonword repetition and phonological awareness). These phenotypes were investigated for linkage with markers on chromosomes 1p36, 6p22, and 15q21; all regions had been previously reported as susceptibility regions for developmental dyslexia. The results of this linkage investigation indicated linkage of the speech and nonword reading phenotypes with the regions on chromosomes 6 and 15, with suggestive results for the chromosome 1 locus.

Concluding Remarks

A review of the evidence presented here permits an unequivocal conclusion that the development and manifestation of DSL is driven, at least partially and substantially, by the variation in the human genome (See Text Box 2). Yet, in addition to this general statement, only a few remarks can be made with any degree of certainty.

First, it appears that the genetic mechanisms involved in DSL are complex. At the beginning, the issue of “general” versus “specific” mechanisms unifying or differentiating typical and atypical language acquisition was discussed. In fact, the review of the literature carried out here suggests that this differentiation might be inadequate, because both types of mechanisms might be at work. The example of the KE family indicates that there are might be major genes that differentiate families like KE and other individuals from the general population by the impact of a gene such as *FOXP2*. Given the multidimensional nature of their phenotypes, KE family members and such individuals truly form a subgroup different from the general population—rare distinct examples of DSL. Yet, the mutations in *FOXP2* do not appear to generalize to “common” forms of DSL and account for only a very small number of cases. Thus, it is plausible and likely that other genes, which may or may not be functionally related to *FOXP2*, account for common DSL. Correspondingly, it is unlikely that there are going to be specific genes whose function would be restricted to form the genetic bases for speaking and language acquisition. It is much more likely that there are many genes that contribute to a variety of functions and that these genes form networks that are recruited in the process of forming language-related representations during language acquisition. Depending on the nature and type of particular insults, these networks might flexibly recover their functions or exhibit deficient functioning resulting in the formation of inadequate representational systems. In other words, it is anticipated that the genes for speaking are many, and they are intertwined in different functional networks which pleiotropically support multiple functions.

Second, the landscape of findings discussed here underscores even more the importance of carefully defining the phenotype to be studied. For example, many if not all theories of DSL include deficiencies in phonological representation as an element of the clinical and psychological profile of DSL. Latent deficits in phonological representation can manifest in multiple constructs, which in turn can be captured by observed variables. One such construct is phonological memory, a type of memory that ensures storage and processing of phonemes and their sequences. Phonological memory can be captured and quantified through a number

of processes such as, as discussed earlier, the processing of nonwords—the skill of repeating orally presented nonsense words of varying levels of difficulty. Correspondingly, familiarity and heritability estimates can be obtained for such a skill. For example, in one twin study the estimate of heritability for the indicator of nonword repetition was close to 1.0, suggesting that genes are an important source of individual differences for this trait (Bishop, North, & Donlan, 1996). Based on this and other findings, the phenotype of nonword repetition has been used as a phenotype in molecular-genetic research, and linkage was established to this phenotype (The SLI Consortium, 2004). In addition, the nonword repetition task appears to be the single phenotype that best differentiates affected and unaffected individuals (Watkins et al., 2002). It is also of interest that, when samples of parents of probands with SLI and typically developing children were compared on a number of psychological indicators relevant to linguistic functioning (Barry, Yasin, & Bishop, 2006), the two groups differed significantly in their performance on nonword repetition, oral-motor, and digit span tasks. Although all three tasks highlighted important information that differentiated the two groups, the nonword repetition task alone discriminated 75.8% of the cases. The authors concluded that nonword repetition can be viewed as a reliable marker of family risk for language impairment. Finally, it has been shown in longitudinal studies of SLI that twins with SLI had deficits in nonword repetition even when the holistic, syndromatic presentation of the disorder was not present (Bishop et al., 1996). Therefore, it is possible that there are particular “dimensions” of all DSL that capture deficits common to all of them. Yet, given the diversity in the presentation and course of different DSL, it is likely that each disorder will have its unique dimensions.

Third, this “partial overlap/partial specificity” of multidimensional phenotypes of different DSL might lead to what is now referred to as the retrospective and prospective comorbidity of these disorders. For example, the typically observed onset of stuttering is between the ages of 3 and 6 years, and reports rates of natural, unassisted recovery of ~75% (Yairi & Ambrose, 1999, 2005). Thus, the prevalence of stuttering as a lifetime disorder is much lower than its incidence [.5 to 1% vs. 4 to 5%, respectively, (Bloodstein, 1995; Felsenfeld, 2002)]. Yet, childhood stuttering is a significant risk factor for other DSL that develop later in life, even if stuttering stops manifesting (Ardila et al., 1994). It is possible that this “continuity” of developmental transformation is due to the presence of particular dimensions of the complex phenotypes common to all DSL.

Fourth, the issue of pleiotropy, or the impact of the same genes on multiple phenotypes, has also been discussed in the literature on DSL, given the substantive overlap in regions of linkage for a variety of developmental disorders such as SSD and developmental dyslexia (Miscimarra et al., 2007; Stein et al., 2006; Stein et al., 2004), and SLI and autism (Tager-Flusberg & Joseph, 2003). Once again, whether these are true examples of pleiotropy or outcomes of the imprecision of phenotype definitions is yet to be determined. In this context, molecular-genetic studies of DSL can be instrumental in re-categorizing probands into those having particular genetic defects (e.g., particular mutations) and those not having such defects and exploring their presentational features. Such classifications of probands based on their genetic etiology might result in the establishment of more precise phenotypes and, possibly, the re-design of clinical classifications. Nevertheless, although the field of behavior- and molecular-genetic studies of DSL today presents a variegated quilt of findings

with only a limited number of common threads, it has been enormously productive in generating exciting findings and providing leads and examples for studies of the genetic etiology of complex human behaviors.

Finally, given today's atmosphere of translational research, it is important to comment on the futuristic but quickly approaching applications of the findings presented above. The most pronounced expectations for genetic and genomic sciences with regard to DSL pertain to diagnosis and treatment. The underlying drive for the discovery of the genetic bases of DSL is in identifying genetic markers that can enhance diagnostic paradigms and in discovering genetic pathways related to DSL, so that pharmacological interventions can be developed. Both possibilities are still in the future, but they re-emphasize the meaning and value of the efforts described above, not only for understanding the etiology of DSL but also for translating this understanding into the diagnosis and treatment of children and adults suffering from these disorders.

Acknowledgments

Preparation of this article was supported by Grants R21 TW006764-02 from the Fogarty Program as administered by the National Institutes of Health, Department of Health and Human Services (PI: Grigorenko), and R01 DC007665 as administered by the National Institute of Deafness and Communication Disorders (PI: Grigorenko). Grantees undertaking such projects are encouraged to express their professional judgment freely. Therefore, this article does not necessarily reflect the position or policies of the National Institutes of Health, and no official endorsement should be inferred. I am thankful to Ms. Kelly Nedwick and Ms. Mei Tan for their editorial assistance and to Susan Felsenfeld for her helpful comments.

References

- Ardila A, Bateman JR, Nino CR, Pulido E, Rivera DB, Vanegas CJ. An epidemiologic study of stuttering. *Journal of Communication Disorders*. 1994; 27:37–48. [PubMed: 8006205]
- Barry JG, Yasin I, Bishop DV. Heritable risk factors associated with language impairments. *Genes, Brain & Behavior*. 2006; 6:66–76.
- Bartlett CW, Flax JF, Logue MW, Smith BJ, Vieland VJ, Tallal P, et al. Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Human Heredity*. 2004; 57:10–20. [PubMed: 15133308]
- Bartlett CW, Flax JF, Logue MW, Vieland VJ, Bassett AS, Tallal P, et al. A major susceptibility locus for specific language impairment is located on 13q21. *American Journal of Human Genetics*. 2002; 71:45–55. [PubMed: 12048648]
- Bates EA. Explaining and interpreting deficits in language development across clinical groups: where do we go from here? *Brain & Language*. 2004; 88:248–253. [PubMed: 14965545]
- Belton E, Salmond CH, Watkins KE, Vargha-Khadem F, Gadian DG. Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Human Brain Mapping*. 2003; 18:194–200. [PubMed: 12599277]
- Bishop DV. Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics*. 1994a; 15:507–550.
- Bishop DV. Is specific language impairment a valid diagnostic category? Genetic and psycholinguistic evidence. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*. 1994b; 346:105–111.
- Bishop DV. Developmental cognitive genetics: how psychology can inform genetics and vice versa. *Quarterly Journal of Experimental Psychology*. 2005; 59:1153–1168.

- Bishop DV, Adams CV, Norbury CF. Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain & Behavior*. 2006; 5:158–169.
- Bishop DV, Laws G, Adams C, Norbury CF. High heritability of speech and language impairments in 6-year-old twins demonstrated using parent and teacher report. *Behavior Genetics*. 2006; 36:173–184. [PubMed: 16485179]
- Bishop DV, North T, Donlan C. Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *Journal of Child Psychology & Psychiatry & Allied Disciplines*. 1996; 37:391–403.
- Bloodstein, O. A handbook on stuttering. Chicago, IL: National Easter Seal Society; 1995.
- Choudhury N, Benasich AA. A family aggregation study: the influence of family history and other risk factors on language development. *Journal of Speech Language & Hearing Research*. 2003; 46:261–272.
- Cole KN, Schwartz IS, Notari AR, Dale PS, Mills PE. Examination of the stability of two methods of defining specific language impairment. *Applied Psycholinguistics*. 1995; 16:103–123.
- DeThorne LS, Hart SA, Petrill SA, Deater-Deckard K, Thompson LA, Schatschneider C, et al. Children's history of speech-language difficulties: genetic influences and associations with reading-related measures. *Journal of Speech Language & Hearing Research*. 2006; 49:1280–1293.
- Felsenfeld S. Finding susceptibility genes for developmental disorders of speech: the long and winding road. *Journal of Communication Disorders*. 2002; 35:329–345. [PubMed: 12160352]
- Felsenfeld S, Plomin R. Epidemiological and offspring analyses of developmental speech disorders using data from the Colorado Adoption Project. *Journal of Speech, Language and Hearing Research*. 1997; 41:778–791.
- Fisher SE, Lai CS, Monaco AP. Deciphering the genetic basis of speech and language disorders. *Annual Review of Neuroscience*. 2003; 26:57–80.
- French CA, Groszer M, Preece C, Coupe AM, Rajewsky K, Fisher SE. Generation of mice with a conditional *Foxp2* null allele. *Genesis: the Journal of Genetics & Development*. 2007; 45:440–446.
- Haines J, Camarata S. Examination of candidate genes in language disorder: a model of genetic association for treatment studies. *Mental Retardation & Developmental Disabilities Research Reviews*. 2004; 10:208–217. [PubMed: 15611983]
- Hurst JA, Baraitser M, Auger E, Graham F, Norell S. An extended family with a dominantly inherited speech disorder. *Developmental Medicine and Child Neurology*. 1990; 32:352–355. [PubMed: 2332125]
- Joanisse MF. Specific Language Impairments in children phonology, semantics, and the English past tense. *Current Directions in Psychological Science*. 2004; 13:156–160.
- Kaufmann E, Knöchel W. Five years on the wings of fork head. *Mechanisms of Development*. 1996; 57:3–20. [PubMed: 8817449]
- Kovas Y, Hayiou-Thomas ME, Oliver B, Dale PS, Bishop DV, Plomin R. Genetic influences in different aspects of language development: The etiology of language skills in 4.5-year-old twins. *Child Development*. 2005; 76:632–651. [PubMed: 15892783]
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 2001; 413:519–523. [PubMed: 11586359]
- Leonard, LB. Children with specific language impairment. Cambridge, MA: MIT Press; 1998.
- Levis B, Ricci D, Lukong J, Drayna D. Genetic linkage studies in a large West African kindred. *American Journal of Human Genetics*. 2004; 75:S2026.
- MacDermot KD, Bonora E, Sykes N, Coupe AM, Lai CS, Vernes SC, et al. Identification of *FOXP2* truncation as a novel cause of developmental speech and language deficits. *American Journal of Human Genetics*. 2005; 76:1074–1080. [PubMed: 15877281]
- Meaburn E, Dale PS, Craig IW, Plomin R. Language-impaired children: No sign of the *FOXP2* mutation. *Cognitive Neuroscience and Neuropsychology*. 2002; 13:1075–1077.
- Miscimarra LE, Stein CM, Millard C, Kluge A, Cartier KC, Freebairn LA, et al. Further evidence of pleiotropy influencing speech and language: Analysis of the *DYX8* region. *Human Heredity*. 2007; 63:47–58. [PubMed: 17230025]

- Monaco AP. The SLI Consortium (SLIC). Multivariate linkage analysis of specific language impairment (SLI). *Annals of Human Genetics*. 2007; 71:660–673. [PubMed: 17388790]
- Newbury DF, Bonora E, Lamb JA, Fisher SE, Lai CSL, Baird G, et al. FOXP2 is not a major susceptibility gene for autism or specific language impairment. *American Journal of Human Genetics*. 2002; 70:1318–1327. [PubMed: 11894222]
- O'Brien EK, Zhang X, Nishimura C, Tomblin JB, Murray JC. Association of Specific Language Impairment (SLI) to the region of 7q31. *American Journal of Human Genetics*. 2003; 72:1536–1543. [PubMed: 12721956]
- Oliver BR, Plomin R. Twins' Early Development Study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behavior problems from childhood through adolescence. *Twin Research & Human Genetics: the Official Journal of the International Society for Twin Studies*. 2007; 10:96–105. [PubMed: 17539369]
- Pinker, S. Words and rules: The ingredients of language. New York, NY, US: Basic Books; 1999.
- Riaz N, Steinberg S, Ahmad J, Pluzhnikov A, Riazuddin S, Cox NJ, et al. Genomewide significant linkage to stuttering on chromosome 12. *American Journal of Human Genetics*. 2005; 76:647–651. [PubMed: 15714404]
- Rice ML, Wexler K, Cleave PL. Specific language impairment as a period of extended optional infinitive. *Journal of Speech & Hearing Research*. 1995; 38:850–863. [PubMed: 7474978]
- Shriberg, LD. Classification and misclassification of child speech sound disorders; Paper presented at the The Annual Conference of the American Speech-Language-Hearing Association; Atlanta. 2002 Nov 22-24.
- Shugart YY, Mundorff J, Kilshaw J, Doheny K, Doan B, Wanyee J, et al. Results of a genome-wide linkage scan for stuttering. *American Journal of Medical Genetics*. 2004; 124:133–135. [PubMed: 14699610]
- Smith SD. Genes, language development, and language disorders. *Mental Retardation & Developmental Disabilities Research Reviews*. 2007; 13:96–105. [PubMed: 17326114]
- Smith SD, Pennington BF, Boada R, Shriberg LD. Linkage of speech sound disorder to reading disability loci. *Journal of Child Psychology and Psychiatry*. 2005; 46:1057–1066. [PubMed: 16178929]
- Stein CM, Millard C, Kluge A, Miscimarra LE, Cartier KC, Freebairn LA, et al. Speech sound disorder influenced by a locus in 15q14 region. *Behavior Genetics*. 2006; 36:858–868. [PubMed: 16786424]
- Stein CM, Schick JH, Taylor GH, Shriberg LD, Millard C, Kundtz-Kluge A, et al. Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. *American Journal of Human Genetics*. 2004; 74:283–297. [PubMed: 14740317]
- Stromswold K. Genetics of spoken language disorders. *Human Biology*. 1998; 70:297, 291–212. [PubMed: 9549241]
- Stromswold K. The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language*. 2001; 77:647–723.
- Suresh R, Ambrose N, Roe C, Pluzhnikov A, Wittke-Thompson JK, Ng MCY, et al. New complexities in the genetics of stuttering: Significant sex-specific linkage signals. *American Journal of Human Genetics*. 2006; 78:554–563. [PubMed: 16532387]
- Tager-Flusberg H, Joseph RM. Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*. 2003; 358:303–314.
- The SLI Consortium. A genomewide scan identifies two novel loci involved in SLI. *American Journal of Human Genetics*. 2002; 70:384–398. [PubMed: 11791209]
- The SLI Consortium. Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by SLI. *American Journal of Human Genetics*. 2004; 74:1225–1238. [PubMed: 15133743]
- Verhoeven, L.; van Balkom, H., editors. Classification of developmental language disorders. Mahwah, NJ: Lawrence Erlbaum; 2004.
- Watkins KE, Dronkers NF, Vargha-Khadem F. Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*. 2002; 125:452–464. [PubMed: 11872604]

- White SA, Fisher SE, Geschwind DH, Scharff C, Holy TE. Singing mice, songbirds, and more: models for FOXP2 function and dysfunction in human speech and language. *Journal of Neuroscience*. 2006; 26:10376–10379. [PubMed: 17035521]
- Wittke-Thompson JK, Ambrose N, Yairi E, Roec C, Cook EH, Ober C, et al. Genetic studies of stuttering in a founder population. *Journal of Fluency Disorders*. 2007; 32:33–50. [PubMed: 17276504]
- Yairi E, Ambrose NG. Early childhood stuttering I: Persistency and recovery rates. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1097–1112.
- Yairi, E.; Ambrose, NG. Early childhood stuttering. Austin, TX: Pro-Ed; 2005.

Abbreviations

SSD	Speech and Sound Disorders
SLI	Specific Language Impairment
DSL	Disorders of Speech and Language

Text Box (1)**Types of psychological representations forming the texture of language-related skills**

The *phonetic* (or *phonological*) representation of a word pertains to its pronunciation and understanding, whether spoken or written. *Phonemes* are the smallest units of sound that form the texture of meaningful speech, whether oral or printed, in a given language.

Every word in a language is characterized by a particular internal structure. *Morphology* pertains to studies of word structure or word formation and relationships between words; correspondingly, *morphological* and *lexical* representations capture this type of word-related information.

When words are connected together in a manner specific to particular languages, they form sentences. Internal structures of sentences and rules of word connectedness are captured by *syntactic* representations.

Every real word in a language has a meaning or several meanings associated with it. This aspect of words and languages is captured by *semantic* representations.

The knowledge of the appropriate use of words, phrases, and sentences in the context of discourse or conversation is captured by *pragmatic* representation.

Text Box (2)**Summative Observations on the Etiology of DLS**

1. Disorders of Speech and Language are familial and heritable.
2. These disorders characterize an atypical acquisition of various types of psychological representations (e.g., phonological, morphological, lexical, semantic, syntactic, and pragmatic), each of which is also familial and heritable both in general and disordered populations.
3. Specific manifestations of Disorders of Speech and Language can be transmitted in simple Mendelian fashion, but, in such cases, they capture only very rare manifestations of such disorders which can be associated with disruptions in specific single genes (e.g., a complex speech and language disturbance in the KE family, which is associated with a single point mutation in the *FOXP2* gene).
4. Common forms of Disorders of Speech and Language are mostly likely associated with variability in the function of multiple genes.
5. The field has revealed a number of interesting candidate regions for Disorders of Speech and Language, but no specific candidate genes have been identified and sufficiently interrogated to be referred to as a candidate gene for Disorders of Speech and Language.

Table 1

Whole-genome Studies of DSL

Type of participants	Type of markers	Phenotypes	Regions identified	Evidence	Source
Large Canadian families (n=5) of Celtic ancestry	STRPs (at ~9 cM)	(1) language impairment (2) reading impairment	2p22 13q21 17q23	Suggestive (LOD=2.86) Significant (LOD=3.92) Suggestive (LOD=2.19)	(Bartlett et al., 2002)
Small UK nuclear families (n=98)/sibpairs (n=219)	STRPs (at ~8 cM)	(3) clinical impairment	16q24 19q13 18q	Significant (LOD=3.55) Significant (LOD=3.55) (LOD=1.51)	(The SLI Consortium, 2002) (Shugart et al., 2004)
Small nuclear families (n=68) of European ancestry	STRPs (at ~10 cM)	Nonword repetition	1q21-1q22	Suggestive (LOD=2.27)	(Levis, Ricci, Lukong, & Drayna, 2004)
A large Cameroonian family	STRPs (at ~10 cM)	Expressive language	12q	Significant (LOD=4.61)	(Riaz et al., 2005)
Highly inbred Pakistani families (n =46)	STRPs (at ~10.5 cM)	Stuttering	1q	Suggestive (LOD=2.3)	(Suresh et al., 2006)
Nuclear families of European descent (n=100) from The Illinois International Genetics of Stuttering Project	10,000 SNPs (at ~1 SNP per 100 kb)	Stuttering	9q 15q	(LOD=1.95)	
A large Hutterite family	1,200 STRPs and SNPs	Stuttering	7q (males only) 21q (females only) 3q 13q 15q	Suggestive (LOD=2.99) Significant (LOD=4.5) (p=.013) (p=.012) (p=.020)	(Witke-Thompson et al., 2007)

Note: STRP=Short-tandem repeat polymorphism; SNP=single-nucleotide polymorphism