The Genetics of Speech and Language Impairments
Karin Stromswold, M.D., Ph.D.

Without instruction, most children master the complexities of spoken language by the age of 6 or 7 years. About 5% of apparently healthy children, however, struggle to acquire basic competence in one or more aspects of spoken language and are classified as having specific language impairment.

Genetic factors have an important role in many such cases. Children with specific language impairment are four times as likely to have a family history of the disorder as are children who do not have such an impairment, and the concordance rate for the disorder is almost twice as great for monozygotic twins as for dizygotic twins. More than 10 susceptibility loci have been identified. More often than not, loci that are robustly linked to specific language impairment in one study show no linkage in other studies, and all these loci have been linked to other neurodevelopmental disorders. Are these reported associations real? If so, which genes underlie these linkages, and what is their mechanistic effect?

Perhaps one reason that linkage studies have implicated different loci for specific language impairment is that each group of investigators has used different case definitions. Since the subjects are selected in different ways and different measures are used to define language impairment, the discovery of different loci would not be unexpected. Studies of relatively homogeneous groups of children with specific language impairment, whose disorders would appear to have a common cause, would seem to be more likely to yield a robust genetic result. Causes that have been proposed for receptive specific language impairment include deficits in short-term auditory memory, auditory sequencing, and rapid auditory processing.

A second way of dealing with the clinical heterogeneity is to define a measurable biologic marker, or endophenotype, that is present in children with specific language impairment regardless of how the disorder is manifested clinically. However, even carefully delineated endophenotypes may have different causes. For example, children might perform poorly in the repetition of nonsense words because they have poor auditory abilities related to short-term memory, sequencing, processing, or perception or because they have oral motor apraxia or are simply inattentive during testing. Moreover, the same genotype can result in different endophenotypes. A person carrying a mutation that affects the coordination of complex oral motor movements could present as someone who performs poorly when asked to repeat nonsense words, is mute or selectively mute, omits phonologically unstressed elements and hence appears to have a grammatical deficit, or has speech dyspraxia, dysfluency, or stutter.

A third and powerful technique for elucidating the genetics of specific language impairment is to identify cases of language impairments that show mendelian transmission. Virtually all cases of familial language impairments have multifactorial, polygenic transmission. However, in 1990, Hurst et al. discovered a three-generation British family with autosomal dominant transmission of oral motor and speech dyspraxia. AFFECTED members of this family (called the KE family) carry a mutation in FOXP2, which encodes a transcription factor. However, it is not clear how the mutation results in speech dyspraxia in affected
family members. Furthermore, the extent to which this mutation is specific to speech is unclear, because affected family members also have oral motor dyspraxia, low nonverbal IQs, and nonverbal learning disorders.7

The FOXP2 protein has many transcription targets in the brain.9,10 In this issue of the Journal, Vernes et al.11 report that FOXP2 down-regulates the expression of CNTNAP2, a gene that encodes a neurexin protein. Using scores on tests of nonsense-word repetition for children with specific language impairment as an endophenotype, they further show that particular CNTNAP2 variants are associated with specific language impairment.

This study represents a bold step in elucidation of the genetic underpinnings of language impairment. One wonders, though, why large studies have not uncovered linkage between specific language impairment and 7q35–q36, the chromosomal locus of CNTNAP2. More specifically, in previous linkage studies,1,2 which were carried out using the same group of children analyzed by Vernes et al., why did investigators not report even a hint of linkage between scores on nonsense-word repetition and the 7q35–q36 locus? Why did the authors choose CNTNAP2 as the sole target gene to investigate rather than selecting among the more than four dozen candidate genes they previously identified in their scans for FOXP2-regulated genes?9,10 The previous studies identified candidate genes that are selectively expressed in the frontal cortex (which is associated with language) and the basal ganglia (which have been implicated in apraxia), and both of these areas are anomalous in affected KE family members.7,8 Why, for example, did Vernes et al. not select SLC17A3, a gene expressed in basal ganglia and identified as a candidate in both of these studies?9,10 Why not one of the WNT genes (identified in one of the studies), which is expressed in human brain development and has been associated with autism, schizophrenia, and Alzheimer’s disease?12 Finally, why did they choose scores on nonsense-word repetition as the sole endophenotype for specific language impairment, given that the probands were identified with the use of another test?

The general relevance of CNTNAP2 to speech dyspraxia remains to be determined. As noted by Vernes et al., variants of CNTNAP2 have been associated with neurodevelopmental disorders whose primary manifestation is not linguistic. Particularly intriguing are three studies showing association between CNTNAP2 variants and autism. Autistic children often have deficits in nonsense-word repetition,12-14 a finding consistent with a role of CNTNAP2 in nonsense-word repetition. Perhaps CNTNAP2 will be implicated in other clinical populations with disorders involving nonsense-word repetition.

According to the common variant–multiple disease hypothesis,15 common alleles that contribute to a particular disease under particular genetic and environmental conditions may result in a different disease (or no disease) under other genetic and environmental conditions. For a group of related disorders, some etiologic factors are unique to a particular disease, and other factors are shared by several diseases. This hypothesis explains why loci associated with spoken-language disorders are also linked with other neurodevelopmental disorders, why different people with the same genetic mutation have different clinical pictures, and why linkage analyses of people with familial language disorders often do not identify susceptibility loci, including those previously identified. The developmental language disorders can thus be placed in a larger class of neurodevelopmental disorders. By carefully delineating endophenotypes that map out what is known about the processes of language and language disorders and by carrying out careful studies like that of Vernes et al., we can begin to understand how genetic and environmental factors affect language and language disorders.

No potential conflict of interest relevant to this article was reported.
EDITORIAL


Copyright © 2008 Massachusetts Medical Society.