

Genetic and Psycholinguistic Evidence

CSC2540S Machine Learning and Universal Grammar
Department of Computer Science, University of Toronto

Shalom Lappin*
Department of Philosophy
King's College, London

* Joint work with Alex Clark
Department of Computer Science
Royal Holloway College, London

March 17, 2009

Outline

- 1 An Inherited Language Disorder: The FOXP2 Gene
- 2 Specific Language Impairment
- 3 Bayesian Inference in Human Language Acquisition
- 4 Conclusions

The KE Family

- The KE family is a large extended family, half of whom are affected by an inherited speech and language deficit.
- Hurst et al. (1990) provide the first report on the family, whose affected members they describe as suffering from severe developmental *verbal apraxia*.
- This condition is an oral motor disorder that disrupts the production of speech sounds.
- In addition, affected members exhibit difficulty in producing and processing syllable sequences, morphological paradigms, and well formed syntactic structures, and they suffer from non-linguistic oral praxis.

The KE Family

- The KE family is a large extended family, half of whom are affected by an inherited speech and language deficit.
- Hurst et al. (1990) provide the first report on the family, whose affected members they describe as suffering from severe developmental *verbal apraxia*.
- This condition is an oral motor disorder that disrupts the production of speech sounds.
- In addition, affected members exhibit difficulty in producing and processing syllable sequences, morphological paradigms, and well formed syntactic structures, and they suffer from non-linguistic oral praxis.

The KE Family

- The KE family is a large extended family, half of whom are affected by an inherited speech and language deficit.
- Hurst et al. (1990) provide the first report on the family, whose affected members they describe as suffering from severe developmental *verbal apraxia*.
- This condition is an oral motor disorder that disrupts the production of speech sounds.
- In addition, affected members exhibit difficulty in producing and processing syllable sequences, morphological paradigms, and well formed syntactic structures, and they suffer from non-linguistic oral praxis.

The KE Family

- The KE family is a large extended family, half of whom are affected by an inherited speech and language deficit.
- Hurst et al. (1990) provide the first report on the family, whose affected members they describe as suffering from severe developmental *verbal apraxia*.
- This condition is an oral motor disorder that disrupts the production of speech sounds.
- In addition, affected members exhibit difficulty in producing and processing syllable sequences, morphological paradigms, and well formed syntactic structures, and they suffer from non-linguistic oral praxis.

The Genetic Basis of the Disorder

- Lai et al. (2001) studied three generations of the KE family, and they identify the precise genetic element of the language deficit.
- They show that in members of the family suffering from the disorder the FOXP2 gene is mutated at one point in its amino acid coding sequence.
- The FOXP2 gene encodes an inscription factor for messenger RNA that controls binding in the target DNA of forkhead proteins, which are important regulating elements in embryogenesis.

The Genetic Basis of the Disorder

- Lai et al. (2001) studied three generations of the KE family, and they identify the precise genetic element of the language deficit.
- They show that in members of the family suffering from the disorder the FOXP2 gene is mutated at one point in its amino acid coding sequence.
- The FOXP2 gene encodes an inscription factor for messenger RNA that controls binding in the target DNA of forkhead proteins, which are important regulating elements in embryogenesis.

The Genetic Basis of the Disorder

- Lai et al. (2001) studied three generations of the KE family, and they identify the precise genetic element of the language deficit.
- They show that in members of the family suffering from the disorder the FOXP2 gene is mutated at one point in its amino acid coding sequence.
- The FOXP2 gene encodes an inscription factor for messenger RNA that controls binding in the target DNA of forkhead proteins, which are important regulating elements in embryogenesis.

A Mutation in the FOXP2 Gene

- The mutation of the FOXP2 gene results in the substitution of a single amino acid (arginine to histidine substitution) in its forkhead binding domain.
- The affected members of the KE family have only one mutated FOXP2 gene in the pair of these genes that they inherit (one from each parent).
- This single mutation disrupts early development in part of the brain through the underproduction of forkhead proteins.
- CE, a person who is not related to the KE family and exhibits the same language disorder, also displays this mutation of the FOXP2 gene.

A Mutation in the FOXP2 Gene

- The mutation of the FOXP2 gene results in the substitution of a single amino acid (arginine to histidine substitution) in its forkhead binding domain.
- The affected members of the KE family have only one mutated FOXP2 gene in the pair of these genes that they inherit (one from each parent).
- This single mutation disrupts early development in part of the brain through the underproduction of forkhead proteins.
- CE, a person who is not related to the KE family and exhibits the same language disorder, also displays this mutation of the FOXP2 gene.

A Mutation in the FOXP2 Gene

- The mutation of the FOXP2 gene results in the substitution of a single amino acid (arginine to histidine substitution) in its forkhead binding domain.
- The affected members of the KE family have only one mutated FOXP2 gene in the pair of these genes that they inherit (one from each parent).
- This single mutation disrupts early development in part of the brain through the underproduction of forkhead proteins.
- CE, a person who is not related to the KE family and exhibits the same language disorder, also displays this mutation of the FOXP2 gene.

A Mutation in the FOXP2 Gene

- The mutation of the FOXP2 gene results in the substitution of a single amino acid (arginine to histidine substitution) in its forkhead binding domain.
- The affected members of the KE family have only one mutated FOXP2 gene in the pair of these genes that they inherit (one from each parent).
- This single mutation disrupts early development in part of the brain through the underproduction of forkhead proteins.
- CE, a person who is not related to the KE family and exhibits the same language disorder, also displays this mutation of the FOXP2 gene.

Is the FOXP2 a "Language Gene"?

- Variants of the FOXP2 gene are present in the genotypes of a wide variety of vertebrate species, including chimps, rodents, and birds.
- Jackendoff and Pinker (2005) observe that the human form of the gene is unique.
- They argue that it constitutes a central component in the genetic mechanism for the encoding and development of the language faculty.
- They suggest that the human form of the gene was selected for its role in promoting the emergence of language.

Is the FOXP2 a "Language Gene"?

- Variants of the FOXP2 gene are present in the genotypes of a wide variety of vertebrate species, including chimps, rodents, and birds.
- Jackendoff and Pinker (2005) observe that the human form of the gene is unique.
- They argue that it constitutes a central component in the genetic mechanism for the encoding and development of the language faculty.
- They suggest that the human form of the gene was selected for its role in promoting the emergence of language.

Is the FOXP2 a "Language Gene"?

- Variants of the FOXP2 gene are present in the genotypes of a wide variety of vertebrate species, including chimps, rodents, and birds.
- Jackendoff and Pinker (2005) observe that the human form of the gene is unique.
- They argue that it constitutes a central component in the genetic mechanism for the encoding and development of the language faculty.
- They suggest that the human form of the gene was selected for its role in promoting the emergence of language.

Is the FOXP2 a "Language Gene"?

- Variants of the FOXP2 gene are present in the genotypes of a wide variety of vertebrate species, including chimps, rodents, and birds.
- Jackendoff and Pinker (2005) observe that the human form of the gene is unique.
- They argue that it constitutes a central component in the genetic mechanism for the encoding and development of the language faculty.
- They suggest that the human form of the gene was selected for its role in promoting the emergence of language.

The Phenotype of the KE Family Language Disorder

- Watkins et al. (2002) give a detailed description of the phenotype of the language disorder present in the KE family.
- This account is based on extensive language and cognitive skills testing of affected and unaffected members of the family.
- In addition, Watkins et al. compared both parts of the family to a group of aphasics suffering from similar speech and language deficits.
- The test results were subjected to pair wise and three way statistical analysis for variance to identify the variables that discriminate the groups from each other.

The Phenotype of the KE Family Language Disorder

- Watkins et al. (2002) give a detailed description of the phenotype of the language disorder present in the KE family.
- This account is based on extensive language and cognitive skills testing of affected and unaffected members of the family.
- In addition, Watkins et al. compared both parts of the family to a group of aphasics suffering from similar speech and language deficits.
- The test results were subjected to pair wise and three way statistical analysis for variance to identify the variables that discriminate the groups from each other.

The Phenotype of the KE Family Language Disorder

- Watkins et al. (2002) give a detailed description of the phenotype of the language disorder present in the KE family.
- This account is based on extensive language and cognitive skills testing of affected and unaffected members of the family.
- In addition, Watkins et al. compared both parts of the family to a group of aphasics suffering from similar speech and language deficits.
- The test results were subjected to pair wise and three way statistical analysis for variance to identify the variables that discriminate the groups from each other.

The Phenotype of the KE Family Language Disorder

- Watkins et al. (2002) give a detailed description of the phenotype of the language disorder present in the KE family.
- This account is based on extensive language and cognitive skills testing of affected and unaffected members of the family.
- In addition, Watkins et al. compared both parts of the family to a group of aphasics suffering from similar speech and language deficits.
- The test results were subjected to pair wise and three way statistical analysis for variance to identify the variables that discriminate the groups from each other.

The Phenotype of the KE Family Language Disorder

- The test results show that the affected KE family group had a significantly lower mean non-verbal intelligence (PIQ) score than either the unaffected family or aphasic group (the latter two did not differ significantly from each other).
- Watkins et al. observe that the only non-verbal subtest on which the affected KE family group suffers from a deficit relative to the unaffected family group is the coding test.
- This task involves associating a series of symbols with a sequence of digits.

The Phenotype of the KE Family Language Disorder

- The test results show that the affected KE family group had a significantly lower mean non-verbal intelligence (PIQ) score than either the unaffected family or aphasic group (the latter two did not differ significantly from each other).
- Watkins et al. observe that the only non-verbal subtest on which the affected KE family group suffers from a deficit relative to the unaffected family group is the coding test.
- This task involves associating a series of symbols with a sequence of digits.

The Phenotype of the KE Family Language Disorder

- The test results show that the affected KE family group had a significantly lower mean non-verbal intelligence (PIQ) score than either the unaffected family or aphasic group (the latter two did not differ significantly from each other).
- Watkins et al. observe that the only non-verbal subtest on which the affected KE family group suffers from a deficit relative to the unaffected family group is the coding test.
- This task involves associating a series of symbols with a sequence of digits.

The Core Deficit

- Watkins et al. suggest that the coding test result might indicate a core deficit in sequencing or in learning associations that underlies both verbal and non-verbal aspects of the disorder.
- They also state

"We cannot rule out the possibility, however, that the genetic abnormality in the KE family produces a general, but mild developmental decay affecting both verbal and non-verbal abilities, as well as a more specific verbal impairment that arises from the articulation deficit." (p.462)

The Core Deficit

- Watkins et al. suggest that the coding test result might indicate a core deficit in sequencing or in learning associations that underlies both verbal and non-verbal aspects of the disorder.
- They also state

"We cannot rule out the possibility, however, that the genetic abnormality in the KE family produces a general, but mild developmental decay affecting both verbal and non-verbal abilities, as well as a more specific verbal impairment that arises from the articulation deficit." (p.462)

Cross Species Comparisons

- White et al. (2006), Groszer et al. (2008), and Fisher (2008) report experiments on mice and finches which involve inducing the same type of point mutation in the respective species counterparts of the human FOXP2 gene that is observed in affected members of the KE family.
- For mice this resulted in impaired motor skill learning, as measured by running on a tilted wheel.
- In the case of finches, the birds with the gene mutation were unable to learn songs from an adult finch.

Cross Species Comparisons

- White et al. (2006), Groszer et al. (2008), and Fisher (2008) report experiments on mice and finches which involve inducing the same type of point mutation in the respective species counterparts of the human FOXP2 gene that is observed in affected members of the KE family.
- For mice this resulted in impaired motor skill learning, as measured by running on a tilted wheel.
- In the case of finches, the birds with the gene mutation were unable to learn songs from an adult finch.

Cross Species Comparisons

- White et al. (2006), Groszer et al. (2008), and Fisher (2008) report experiments on mice and finches which involve inducing the same type of point mutation in the respective species counterparts of the human FOXP2 gene that is observed in affected members of the KE family.
- For mice this resulted in impaired motor skill learning, as measured by running on a tilted wheel.
- In the case of finches, the birds with the gene mutation were unable to learn songs from an adult finch.

The FOXP2 Gene and Synaptic Connections

- Groszer et al. (2008) and Fisher (2008) describe work that shows FOXP2 genes to be essential in the development of synaptic connections in the brain that support motor learning.
- When the point mutation is induced in the counterpart gene in mice, synaptic connections lose their plasticity, which impairs motor learning.
- It is reasonable to hypothesize that the mutation in the human form of FOXP2 has a similar effect.

The FOXP2 Gene and Synaptic Connections

- Groszer et al. (2008) and Fisher (2008) describe work that shows FOXP2 genes to be essential in the development of synaptic connections in the brain that support motor learning.
- When the point mutation is induced in the counterpart gene in mice, synaptic connections lose their plasticity, which impairs motor learning.
- It is reasonable to hypothesize that the mutation in the human form of FOXP2 has a similar effect.

The FOXP2 Gene and Synaptic Connections

- Groszer et al. (2008) and Fisher (2008) describe work that shows FOXP2 genes to be essential in the development of synaptic connections in the brain that support motor learning.
- When the point mutation is induced in the counterpart gene in mice, synaptic connections lose their plasticity, which impairs motor learning.
- It is reasonable to hypothesize that the mutation in the human form of FOXP2 has a similar effect.

A Developmental Motor Learning Disorder

- If this is the case, then the KE family speech and language deficit is caused by a developmental disorder that undermines motor learning and development, with concentration in areas of the brain responsible for articulation and sequence learning.
- This view of the deficit does not require genetic coding of a language faculty or evolutionary selection of the FOXP2 gene for language.

A Developmental Motor Learning Disorder

- If this is the case, then the KE family speech and language deficit is caused by a developmental disorder that undermines motor learning and development, with concentration in areas of the brain responsible for articulation and sequence learning.
- This view of the deficit does not require genetic coding of a language faculty or evolutionary selection of the FOXP2 gene for language.

Genes and Cognition: Against the "Grammar Gene"

- Fisher (2006) argues that the notion of a "language gene" is misguided.
- He points out that the connection between genetic coding and cognitive functions is far too indirect and complex to be accommodated within the view that single genes or gene sequences generate a repertoire like language.
- Genes control development by regulating the production of proteins in cells.
- It is not possible to reduce cognitive capacities to genetic coding sequences.

Genes and Cognition: Against the "Grammar Gene"

- Fisher (2006) argues that the notion of a "language gene" is misguided.
- He points out that the connection between genetic coding and cognitive functions is far too indirect and complex to be accommodated within the view that single genes or gene sequences generate a repertoire like language.
- Genes control development by regulating the production of proteins in cells.
- It is not possible to reduce cognitive capacities to genetic coding sequences.

Genes and Cognition: Against the "Grammar Gene"

- Fisher (2006) argues that the notion of a "language gene" is misguided.
- He points out that the connection between genetic coding and cognitive functions is far too indirect and complex to be accommodated within the view that single genes or gene sequences generate a repertoire like language.
- Genes control development by regulating the production of proteins in cells.
- It is not possible to reduce cognitive capacities to genetic coding sequences.

Genes and Cognition: Against the "Grammar Gene"

- Fisher (2006) argues that the notion of a "language gene" is misguided.
- He points out that the connection between genetic coding and cognitive functions is far too indirect and complex to be accommodated within the view that single genes or gene sequences generate a repertoire like language.
- Genes control development by regulating the production of proteins in cells.
- It is not possible to reduce cognitive capacities to genetic coding sequences.

Grammatical SLI

- van der Lely ((2004), (2005)) argues that there is a sub-type of Specific Language Impairment (SLI) which affects grammatical development (G-SLI).
- She regards G-SLI as independent of other language impairments, in particular, the phonological short term memory (STM) deficit often associated with SLI.
- She claims that G-SLI provides evidence for the existence of a domain specific cognitive module devoted to syntactic representation.
- van der Lely (2004) identifies G-SLI with a developmental disorder that affects a child's grammatical competence.

Grammatical SLI

- van der Lely ((2004), (2005)) argues that there is a sub-type of Specific Language Impairment (SLI) which affects grammatical development (G-SLI).
- She regards G-SLI as independent of other language impairments, in particular, the phonological short term memory (STM) deficit often associated with SLI.
- She claims that G-SLI provides evidence for the existence of a domain specific cognitive module devoted to syntactic representation.
- van der Lely (2004) identifies G-SLI with a developmental disorder that affects a child's grammatical competence.

Grammatical SLI

- van der Lely ((2004), (2005)) argues that there is a sub-type of Specific Language Impairment (SLI) which affects grammatical development (G-SLI).
- She regards G-SLI as independent of other language impairments, in particular, the phonological short term memory (STM) deficit often associated with SLI.
- She claims that G-SLI provides evidence for the existence of a domain specific cognitive module devoted to syntactic representation.
- van der Lely (2004) identifies G-SLI with a developmental disorder that affects a child's grammatical competence.

Grammatical SLI

- van der Lely ((2004), (2005)) argues that there is a sub-type of Specific Language Impairment (SLI) which affects grammatical development (G-SLI).
- She regards G-SLI as independent of other language impairments, in particular, the phonological short term memory (STM) deficit often associated with SLI.
- She claims that G-SLI provides evidence for the existence of a domain specific cognitive module devoted to syntactic representation.
- van der Lely (2004) identifies G-SLI with a developmental disorder that affects a child's grammatical competence.

Grammatical SLI

- van der Lely claims that this disorder results in an inability to deal with complex syntactic structures.
- She characterizes syntactic complexity in terms of dependency relations in a hierarchical constituent phrase structure.
- She analyzes problems with wh-movement exhibited by some SLI test subjects as the misidentification of syntactic movement as an optional rule.
- This analysis is formulated in terms of an early version of Chomsky's minimalist program (Chomsky (1995)), in a highly theory dependent way.

Grammatical SLI

- van der Lely claims that this disorder results in an inability to deal with complex syntactic structures.
- She characterizes syntactic complexity in terms of dependency relations in a hierarchical constituent phrase structure.
- She analyzes problems with wh-movement exhibited by some SLI test subjects as the misidentification of syntactic movement as an optional rule.
- This analysis is formulated in terms of an early version of Chomsky's minimalist program (Chomsky (1995)), in a highly theory dependent way.

Grammatical SLI

- van der Lely claims that this disorder results in an inability to deal with complex syntactic structures.
- She characterizes syntactic complexity in terms of dependency relations in a hierarchical constituent phrase structure.
- She analyzes problems with *wh*-movement exhibited by some SLI test subjects as the misidentification of syntactic movement as an optional rule.
- This analysis is formulated in terms of an early version of Chomsky's minimalist program (Chomsky (1995)), in a highly theory dependent way.

Grammatical SLI

- van der Lely claims that this disorder results in an inability to deal with complex syntactic structures.
- She characterizes syntactic complexity in terms of dependency relations in a hierarchical constituent phrase structure.
- She analyzes problems with *wh*-movement exhibited by some SLI test subjects as the misidentification of syntactic movement as an optional rule.
- This analysis is formulated in terms of an early version of Chomsky's minimalist program (Chomsky (1995)), in a highly theory dependent way.

Heritability of Different Types of SLI

- Bishop et al. (2006) tested six year old twins exhibiting SLI for phonological STM, verb inflection (finite and infinitival verb forms), and comprehension of complex syntactic structures.
- They measure the heritability of distinct types of language impairment by determining their statistical independence of environmental factors in identical and non-identical SLI twins.
- They also measure the co-heritability of distinct types of language impairment.

Heritability of Different Types of SLI

- Bishop et al. (2006) tested six year old twins exhibiting SLI for phonological STM, verb inflection (finite and infinitival verb forms), and comprehension of complex syntactic structures.
- They measure the heritability of distinct types of language impairment by determining their statistical independence of environmental factors in identical and non-identical SLI twins.
- They also measure the co-heritability of distinct types of language impairment.

Heritability of Different Types of SLI

- Bishop et al. (2006) tested six year old twins exhibiting SLI for phonological STM, verb inflection (finite and infinitival verb forms), and comprehension of complex syntactic structures.
- They measure the heritability of distinct types of language impairment by determining their statistical independence of environmental factors in identical and non-identical SLI twins.
- They also measure the co-heritability of distinct types of language impairment.

Heritability of Different Types of SLI

- Bishop et al. report that in their sample (163 twin pairs with SLI and a control group of 183 typical language development twin pairs) phonological STM and verb inflection deficits are derived from distinct genetic sources, but their co-occurrence shows modest statistical significance.
- This correlation is also reported in previous work (Conti-Ramsden et.al (2001), Norbury et. al (2001)), and it requires explanation.

Heritability of Different Types of SLI

- Bishop et al. report that in their sample (163 twin pairs with SLI and a control group of 183 typical language development twin pairs) phonological STM and verb inflection deficits are derived from distinct genetic sources, but their co-occurrence shows modest statistical significance.
- This correlation is also reported in previous work (Conti-Ramsden et.al (2001), Norbury et. al (2001)), and it requires explanation.

Heritability of Different Types of SLI

- Bishop et al. also state that while there is "suggestive evidence" for a genetic connection between verb inflection and sentence structure comprehension deficits, this relationship is not statistically significant in their experiments.
- It would be odd if these two grammatical disorders were related to distinct genetic sources.

Heritability of Different Types of SLI

- Bishop et al. also state that while there is "suggestive evidence" for a genetic connection between verb inflection and sentence structure comprehension deficits, this relationship is not statistically significant in their experiments.
- It would be odd if these two grammatical disorders were related to distinct genetic sources.

SLI and Working Memory Deficit

- There is a considerable amount of evidence suggesting a relationship between SLI and dysfunction in working memory.
- Archibald and Gathercole (2006) report a study indicating that children with SLI suffer significant deficits in both verbal STM and working memory.
- Marton et. al (2006) show that in a group of Hungarian speaking SLI children the grammatical deficit was sensitive to morphological complexity, rather than to surface order or length.
- Hungarian, unlike English, is morphologically rich and exhibits relatively free word order without overt syntactic movement.

SLI and Working Memory Deficit

- There is a considerable amount of evidence suggesting a relationship between SLI and dysfunction in working memory.
- Archibald and Gathercole (2006) report a study indicating that children with SLI suffer significant deficits in both verbal STM and working memory.
- Marton et. al (2006) show that in a group of Hungarian speaking SLI children the grammatical deficit was sensitive to morphological complexity, rather than to surface order or length.
- Hungarian, unlike English, is morphologically rich and exhibits relatively free word order without overt syntactic movement.

SLI and Working Memory Deficit

- There is a considerable amount of evidence suggesting a relationship between SLI and dysfunction in working memory.
- Archibald and Gathercole (2006) report a study indicating that children with SLI suffer significant deficits in both verbal STM and working memory.
- Marton et. al (2006) show that in a group of Hungarian speaking SLI children the grammatical deficit was sensitive to morphological complexity, rather than to surface order or length.
- Hungarian, unlike English, is morphologically rich and exhibits relatively free word order without overt syntactic movement.

SLI and Working Memory Deficit

- There is a considerable amount of evidence suggesting a relationship between SLI and dysfunction in working memory.
- Archibald and Gathercole (2006) report a study indicating that children with SLI suffer significant deficits in both verbal STM and working memory.
- Marton et. al (2006) show that in a group of Hungarian speaking SLI children the grammatical deficit was sensitive to morphological complexity, rather than to surface order or length.
- Hungarian, unlike English, is morphologically rich and exhibits relatively free word order without overt syntactic movement.

A Unified Account of distinct SLI Effects

- The effect that Marton et al. report is expected on an account that treats SLI in general, and grammatical impairment in particular, as the result of a deficiency in a common mechanism which undermines a child's capacity to process syntactically and morphologically complex objects.
- It is not clear how van der Lely's claim that problems with syntactic operations like wh-movement are the result of difficulties with a highly localized syntactic module generalizes to Marton's Hungarian SLI data.

A Unified Account of distinct SLI Effects

- The effect that Marton et al. report is expected on an account that treats SLI in general, and grammatical impairment in particular, as the result of a deficiency in a common mechanism which undermines a child's capacity to process syntactically and morphologically complex objects.
- It is not clear how van der Lely's claim that problems with syntactic operations like wh-movement are the result of difficulties with a highly localized syntactic module generalizes to Marton's Hungarian SLI data.

A Unified Account of distinct SLI Effects

- On van der Lely's view it would be necessary to attribute the Hungarian grammatical deficit to a different component of the child's grammatical competence.
- By contrast, taking SLI to be, at least in part, conditioned by a problem in the development of working memory offers the basis for a unified explanation of distinct types of grammatical and phonological impairment.

A Unified Account of distinct SLI Effects

- On van der Lely's view it would be necessary to attribute the Hungarian grammatical deficit to a different component of the child's grammatical competence.
- By contrast, taking SLI to be, at least in part, conditioned by a problem in the development of working memory offers the basis for a unified explanation of distinct types of grammatical and phonological impairment.

Evidence for an Working Memory-Based Explanation

- Marton (2008) describes a study in which Hungarian children with SLI performed below the typical development control group in visual-spatial processing tasks that invoke higher executive functions of working memory, such as attention, filtering of distracting stimuli, and planning.
- However both groups achieved a comparable level in short term visual spatial memory tasks.
- Dodwell and Bavin (2008) report that a group of Australian children with SLI exhibited significant difficulty with inference and the processing of complex information in narratives to which they were exposed.
- They also relate this problem to a deficit in the executive functions of working memory.

Evidence for an Working Memory-Based Explanation

- Marton (2008) describes a study in which Hungarian children with SLI performed below the typical development control group in visual-spatial processing tasks that invoke higher executive functions of working memory, such as attention, filtering of distracting stimuli, and planning.
- However both groups achieved a comparable level in short term visual spatial memory tasks.
- Dodwell and Bavin (2008) report that a group of Australian children with SLI exhibited significant difficulty with inference and the processing of complex information in narratives to which they were exposed.
- They also relate this problem to a deficit in the executive functions of working memory.

Evidence for an Working Memory-Based Explanation

- Marton (2008) describes a study in which Hungarian children with SLI performed below the typical development control group in visual-spatial processing tasks that invoke higher executive functions of working memory, such as attention, filtering of distracting stimuli, and planning.
- However both groups achieved a comparable level in short term visual spatial memory tasks.
- Dodwell and Bavin (2008) report that a group of Australian children with SLI exhibited significant difficulty with inference and the processing of complex information in narratives to which they were exposed.
- They also relate this problem to a deficit in the executive functions of working memory.

Evidence for an Working Memory-Based Explanation

- Marton (2008) describes a study in which Hungarian children with SLI performed below the typical development control group in visual-spatial processing tasks that invoke higher executive functions of working memory, such as attention, filtering of distracting stimuli, and planning.
- However both groups achieved a comparable level in short term visual spatial memory tasks.
- Dodwell and Bavin (2008) report that a group of Australian children with SLI exhibited significant difficulty with inference and the processing of complex information in narratives to which they were exposed.
- They also relate this problem to a deficit in the executive functions of working memory.

SLI and UG

- The claim that G-SLI provides support for the view that there is a domain specific, genetically determined mechanism devoted to the acquisition and representation of syntax is based on equivocal evidence that is not compelling.
- The samples of SLI children tested in the experiments that are invoked to motivate this claim are relatively small, and they represent a very limited range of natural languages.
- A grammar deficit and working memory problems may be dissociable experimentally and in terms of heritability.
- However, the correlation between them appears to be robust across tested languages.

SLI and UG

- The claim that G-SLI provides support for the view that there is a domain specific, genetically determined mechanism devoted to the acquisition and representation of syntax is based on equivocal evidence that is not compelling.
- The samples of SLI children tested in the experiments that are invoked to motivate this claim are relatively small, and they represent a very limited range of natural languages.
- A grammar deficit and working memory problems may be dissociable experimentally and in terms of heritability.
- However, the correlation between them appears to be robust across tested languages.

SLI and UG

- The claim that G-SLI provides support for the view that there is a domain specific, genetically determined mechanism devoted to the acquisition and representation of syntax is based on equivocal evidence that is not compelling.
- The samples of SLI children tested in the experiments that are invoked to motivate this claim are relatively small, and they represent a very limited range of natural languages.
- A grammar deficit and working memory problems may be dissociable experimentally and in terms of heritability.
- However, the correlation between them appears to be robust across tested languages.

SLI and UG

- The claim that G-SLI provides support for the view that there is a domain specific, genetically determined mechanism devoted to the acquisition and representation of syntax is based on equivocal evidence that is not compelling.
- The samples of SLI children tested in the experiments that are invoked to motivate this claim are relatively small, and they represent a very limited range of natural languages.
- A grammar deficit and working memory problems may be dissociable experimentally and in terms of heritability.
- However, the correlation between them appears to be robust across tested languages.

SLI and UG

- Recent work on SLI suggests that deficiencies in executive functions of working memory management that play a central role in non-linguistic cognitive tasks frequently accompany grammar disorders.
- Claims to the effect that SLI in general, and G-SLI in particular indicate the existence of a genetically determined grammar faculty are not supported by the data.
- At this point there does not seem to be sufficient evidence to motivate strong conclusions on the cognitive and genetic foundations of the different subtypes of SLI.

SLI and UG

- Recent work on SLI suggests that deficiencies in executive functions of working memory management that play a central role in non-linguistic cognitive tasks frequently accompany grammar disorders.
- Claims to the effect that SLI in general, and G-SLI in particular indicate the existence of a genetically determined grammar faculty are not supported by the data.
- At this point there does not seem to be sufficient evidence to motivate strong conclusions on the cognitive and genetic foundations of the different subtypes of SLI.

SLI and UG

- Recent work on SLI suggests that deficiencies in executive functions of working memory management that play a central role in non-linguistic cognitive tasks frequently accompany grammar disorders.
- Claims to the effect that SLI in general, and G-SLI in particular indicate the existence of a genetically determined grammar faculty are not supported by the data.
- At this point there does not seem to be sufficient evidence to motivate strong conclusions on the cognitive and genetic foundations of the different subtypes of SLI.

Statistical Learning in Language Acquisition

- Recent experimental work in psycholinguistics provides some evidence that Bayesian inference plays a significant role in language acquisition.
- This research involves constructing artificial languages and testing the effect of transitional probabilities between word and phrase boundaries on learning in both children and adults.
- The transitions between internal word and phrasal constituents is high, while boundaries between constituents separating words and phrases are marked by low probabilities.

Statistical Learning in Language Acquisition

- Recent experimental work in psycholinguistics provides some evidence that Bayesian inference plays a significant role in language acquisition.
- This research involves constructing artificial languages and testing the effect of transitional probabilities between word and phrase boundaries on learning in both children and adults.
- The transitions between internal word and phrasal constituents is high, while boundaries between constituents separating words and phrases are marked by low probabilities.

Statistical Learning in Language Acquisition

- Recent experimental work in psycholinguistics provides some evidence that Bayesian inference plays a significant role in language acquisition.
- This research involves constructing artificial languages and testing the effect of transitional probabilities between word and phrase boundaries on learning in both children and adults.
- The transitions between internal word and phrasal constituents is high, while boundaries between constituents separating words and phrases are marked by low probabilities.

Transitional Probabilities and Word Boundaries

- Saffran et. al (1996) report a set of experiments in which 8-month-old infants learn to identify word boundaries in continuous syllable sequences on the basis of a two minute exposure to training data.
- The words are nonsense terms constructed out of three syllable sequences.
- The transitional probabilities between syllables within a word are maximal (set at 1), while those between syllables crossing word boundaries are low (around 0.33).

Transitional Probabilities and Word Boundaries

- Saffran et. al (1996) report a set of experiments in which 8-month-old infants learn to identify word boundaries in continuous syllable sequences on the basis of a two minute exposure to training data.
- The words are nonsense terms constructed out of three syllable sequences.
- The transitional probabilities between syllables within a word are maximal (set at 1), while those between syllables crossing word boundaries are low (around 0.33).

Transitional Probabilities and Word Boundaries

- Saffran et. al (1996) report a set of experiments in which 8-month-old infants learn to identify word boundaries in continuous syllable sequences on the basis of a two minute exposure to training data.
- The words are nonsense terms constructed out of three syllable sequences.
- The transitional probabilities between syllables within a word are maximal (set at 1), while those between syllables crossing word boundaries are low (around 0.33).

Transitional Probabilities and Word Boundaries

- The transitional probability of a syllable pair XY is computed as the conditional probability $P(Y|X)$ according to its Bayesian MLE condition $P(Y|X) = \frac{c(XY)}{c(X)}$.
- The amount of time that an infant directs his/her gaze at the source of a syllable sequence is taken as evidence for recognition of novelty in the data.
- The infants were able to distinguish familiar words heard in the training samples from novel non-words on the basis of very limited exposure to a word set.
- Saffran et. al conclude that they inferred the word boundaries on the basis of the difference in transitional probabilities between word internal syllable sequences and word external pairs.

Transitional Probabilities and Word Boundaries

- The transitional probability of a syllable pair XY is computed as the conditional probability $P(Y|X)$ according to its Bayesian MLE condition $P(Y|X) = \frac{c(XY)}{c(X)}$.
- The amount of time that an infant directs his/her gaze at the source of a syllable sequence is taken as evidence for recognition of novelty in the data.
- The infants were able to distinguish familiar words heard in the training samples from novel non-words on the basis of very limited exposure to a word set.
- Saffran et. al conclude that they inferred the word boundaries on the basis of the difference in transitional probabilities between word internal syllable sequences and word external pairs.

Transitional Probabilities and Word Boundaries

- The transitional probability of a syllable pair XY is computed as the conditional probability $P(Y|X)$ according to its Bayesian MLE condition $P(Y|X) = \frac{c(XY)}{c(X)}$.
- The amount of time that an infant directs his/her gaze at the source of a syllable sequence is taken as evidence for recognition of novelty in the data.
- The infants were able to distinguish familiar words heard in the training samples from novel non-words on the basis of very limited exposure to a word set.
- Saffran et. al conclude that they inferred the word boundaries on the basis of the difference in transitional probabilities between word internal syllable sequences and word external pairs.

Transitional Probabilities and Word Boundaries

- The transitional probability of a syllable pair XY is computed as the conditional probability $P(Y|X)$ according to its Bayesian MLE condition $P(Y|X) = \frac{c(XY)}{c(X)}$.
- The amount of time that an infant directs his/her gaze at the source of a syllable sequence is taken as evidence for recognition of novelty in the data.
- The infants were able to distinguish familiar words heard in the training samples from novel non-words on the basis of very limited exposure to a word set.
- Saffran et. al conclude that they inferred the word boundaries on the basis of the difference in transitional probabilities between word internal syllable sequences and word external pairs.

An Objection to Saffran et al. (1996)

- Yang (2004) disputes Saffran et al.'s conclusion.
- He describes a word identification experiment on a subset of the CHILDES corpus using transitional syllable probabilities.
- The results of the experiment indicate poor recall and precision for this procedure.
- As Yang observes, this is due to the fact that 85% of the words in his test set are monosyllabic, and so there is no significant distinction between intra-word and inter-word transitional probabilities for most of the terms in this corpus.

An Objection to Saffran et al. (1996)

- Yang (2004) disputes Saffran et al.'s conclusion.
- He describes a word identification experiment on a subset of the CHILDES corpus using transitional syllable probabilities.
- The results of the experiment indicate poor recall and precision for this procedure.
- As Yang observes, this is due to the fact that 85% of the words in his test set are monosyllabic, and so there is no significant distinction between intra-word and inter-word transitional probabilities for most of the terms in this corpus.

An Objection to Saffran et al. (1996)

- Yang (2004) disputes Saffran et al.'s conclusion.
- He describes a word identification experiment on a subset of the CHILDES corpus using transitional syllable probabilities.
- The results of the experiment indicate poor recall and precision for this procedure.
- As Yang observes, this is due to the fact that 85% of the words in his test set are monosyllabic, and so there is no significant distinction between intra-word and inter-word transitional probabilities for most of the terms in this corpus.

An Objection to Saffran et al. (1996)

- Yang (2004) disputes Saffran et al.'s conclusion.
- He describes a word identification experiment on a subset of the CHILDES corpus using transitional syllable probabilities.
- The results of the experiment indicate poor recall and precision for this procedure.
- As Yang observes, this is due to the fact that 85% of the words in his test set are monosyllabic, and so there is no significant distinction between intra-word and inter-word transitional probabilities for most of the terms in this corpus.

Problem with Yang's Argument

- Yang claims that his experiment shows that transitional probability is not an adequate cue for word boundary identification in realistic data of the sort that children receive.
- In fact, this claim is seriously under motivated.
- Child directed speech of the kind that appears in CHILDES does not exhaust the linguistic samples to which children are exposed in their normal environments.
- They generally have access to the full range of multi-syllabic utterances in normal adult speech, even when it is not directed to them.

Problem with Yang's Argument

- Yang claims that his experiment shows that transitional probability is not an adequate cue for word boundary identification in realistic data of the sort that children receive.
- In fact, this claim is seriously under motivated.
- Child directed speech of the kind that appears in CHILDES does not exhaust the linguistic samples to which children are exposed in their normal environments.
- They generally have access to the full range of multi-syllabic utterances in normal adult speech, even when it is not directed to them.

Problem with Yang's Argument

- Yang claims that his experiment shows that transitional probability is not an adequate cue for word boundary identification in realistic data of the sort that children receive.
- In fact, this claim is seriously under motivated.
- Child directed speech of the kind that appears in CHILDES does not exhaust the linguistic samples to which children are exposed in their normal environments.
- They generally have access to the full range of multi-syllabic utterances in normal adult speech, even when it is not directed to them.

Problem with Yang's Argument

- Yang claims that his experiment shows that transitional probability is not an adequate cue for word boundary identification in realistic data of the sort that children receive.
- In fact, this claim is seriously under motivated.
- Child directed speech of the kind that appears in CHILDES does not exhaust the linguistic samples to which children are exposed in their normal environments.
- They generally have access to the full range of multi-syllabic utterances in normal adult speech, even when it is not directed to them.

Problem with Yang's Argument

- There is no reason to exclude this additional data from the range of evidence that children can make use of when computing transitional probabilities for syllable pairs.
- It is not unreasonable to hypothesize that when one takes account of the full range of evidence available to child language learners, a significant correlation between transitional probability patterns and word boundaries in real language data will prove robust.

Problem with Yang's Argument

- There is no reason to exclude this additional data from the range of evidence that children can make use of when computing transitional probabilities for syllable pairs.
- It is not unreasonable to hypothesize that when one takes account of the full range of evidence available to child language learners, a significant correlation between transitional probability patterns and word boundaries in real language data will prove robust.

Using Transitional Probabilities to Identify Phrase Structure

- Thompson and Newport (2007) use transitional probabilities to investigate learning of phrasal boundaries and constituent structure.
- They describe a series of experiments in which English speaking adults are exposed to training sets of samples from simple artificial languages with six word classes.
- Each class contains three words (the word number of some classes is modified for one of the experiments).
- Phrases consist of word pairs where each element of the pair comes from a distinct word class.

Using Transitional Probabilities to Identify Phrase Structure

- Thompson and Newport (2007) use transitional probabilities to investigate learning of phrasal boundaries and constituent structure.
- They describe a series of experiments in which English speaking adults are exposed to training sets of samples from simple artificial languages with six word classes.
- Each class contains three words (the word number of some classes is modified for one of the experiments).
- Phrases consist of word pairs where each element of the pair comes from a distinct word class.

Using Transitional Probabilities to Identify Phrase Structure

- Thompson and Newport (2007) use transitional probabilities to investigate learning of phrasal boundaries and constituent structure.
- They describe a series of experiments in which English speaking adults are exposed to training sets of samples from simple artificial languages with six word classes.
- Each class contains three words (the word number of some classes is modified for one of the experiments).
- Phrases consist of word pairs where each element of the pair comes from a distinct word class.

Using Transitional Probabilities to Identify Phrase Structure

- Thompson and Newport (2007) use transitional probabilities to investigate learning of phrasal boundaries and constituent structure.
- They describe a series of experiments in which English speaking adults are exposed to training sets of samples from simple artificial languages with six word classes.
- Each class contains three words (the word number of some classes is modified for one of the experiments).
- Phrases consist of word pairs where each element of the pair comes from a distinct word class.

The Design of the Experiments

The training sets contain a canonical phrasal pattern of word class sequences, and variations on these patterns involving

- (a) the presence of repeated phrases,
- (b) optional constituents,
- (c) permutations of phrases (moved constituents), and
- (d) variation in the lexical size of two of the four phrase types.

The Design of the Experiments

The training sets contain a canonical phrasal pattern of word class sequences, and variations on these patterns involving

- (a) the presence of repeated phrases,
- (b) optional constituents,
- (c) permutations of phrases (moved constituents), and
- (d) variation in the lexical size of two of the four phrase types.

The Design of the Experiments

The training sets contain a canonical phrasal pattern of word class sequences, and variations on these patterns involving

- (a) the presence of repeated phrases,
- (b) optional constituents,
- (c) permutations of phrases (moved constituents), and
- (d) variation in the lexical size of two of the four phrase types.

The Design of the Experiments

The training sets contain a canonical phrasal pattern of word class sequences, and variations on these patterns involving

- (a) the presence of repeated phrases,
- (b) optional constituents,
- (c) permutations of phrases (moved constituents), and
- (d) variation in the lexical size of two of the four phrase types.

The Design of the Experiments

- Each of the three conditions in (a)-(c) introduces a significant difference in intra-phrasal vs. inter-phrasal transitional probabilities between word classes.
- The former are set at 1, while the latter are lower.
- For each of these four conditions a control group is exposed to a training set in which the conditions do not apply to discrete phrases, but are formulated only for word classes.
- As a result, there is no substantial difference in the transitional probabilities that hold between different word class pairs in the control language.

The Design of the Experiments

- Each of the three conditions in (a)-(c) introduces a significant difference in intra-phrasal vs. inter-phrasal transitional probabilities between word classes.
- The former are set at 1, while the latter are lower.
- For each of these four conditions a control group is exposed to a training set in which the conditions do not apply to discrete phrases, but are formulated only for word classes.
- As a result, there is no substantial difference in the transitional probabilities that hold between different word class pairs in the control language.

The Design of the Experiments

- Each of the three conditions in (a)-(c) introduces a significant difference in intra-phrasal vs. inter-phrasal transitional probabilities between word classes.
- The former are set at 1, while the latter are lower.
- For each of these four conditions a control group is exposed to a training set in which the conditions do not apply to discrete phrases, but are formulated only for word classes.
- As a result, there is no substantial difference in the transitional probabilities that hold between different word class pairs in the control language.

The Design of the Experiments

- Each of the three conditions in (a)-(c) introduces a significant difference in intra-phrasal vs. inter-phrasal transitional probabilities between word classes.
- The former are set at 1, while the latter are lower.
- For each of these four conditions a control group is exposed to a training set in which the conditions do not apply to discrete phrases, but are formulated only for word classes.
- As a result, there is no substantial difference in the transitional probabilities that hold between different word class pairs in the control language.

The Results of the Experiments

- Thompson and Newport (2007) found that for conditions (a)-(c) the experimental group outperformed the control group in learning both sentence and phrasal structure.
- When all four conditions were combined in a single language, the difference between intra-phrasal and inter-phrasal transitions is substantially increased.
- In an experiment with variants of this language type in which the two groups were exposed to a comparatively small set of canonical sentence patterns (5% of the the training set), the experimental subjects achieved far greater success than the control subjects in learning both sentence and phrasal patterns.

The Results of the Experiments

- Thompson and Newport (2007) found that for conditions (a)-(c) the experimental group outperformed the control group in learning both sentence and phrasal structure.
- When all four conditions were combined in a single language, the difference between intra-phrasal and inter-phrasal transitions is substantially increased.
- In an experiment with variants of this language type in which the two groups were exposed to a comparatively small set of canonical sentence patterns (5% of the the training set), the experimental subjects achieved far greater success than the control subjects in learning both sentence and phrasal patterns.

The Results of the Experiments

- Thompson and Newport (2007) found that for conditions (a)-(c) the experimental group outperformed the control group in learning both sentence and phrasal structure.
- When all four conditions were combined in a single language, the difference between intra-phrasal and inter-phrasal transitions is substantially increased.
- In an experiment with variants of this language type in which the two groups were exposed to a comparatively small set of canonical sentence patterns (5% of the the training set), the experimental subjects achieved far greater success than the control subjects in learning both sentence and phrasal patterns.

The Results of the Experiments

- These results indicate that transitional probabilities can provide an important cue for identifying constituent structure from word sequences.
- The experiments provide data only on syntax learning by adults.
- However, when taken together with Saffran et al.'s research on infant identification of word boundaries, they strongly suggest that Bayesian inference of the kind employed in ML methods in NLP plays a significant role in human language acquisition at a variety of levels of morphological and syntactic structure.

The Results of the Experiments

- These results indicate that transitional probabilities can provide an important cue for identifying constituent structure from word sequences.
- The experiments provide data only on syntax learning by adults.
- However, when taken together with Saffran et al.'s research on infant identification of word boundaries, they strongly suggest that Bayesian inference of the kind employed in ML methods in NLP plays a significant role in human language acquisition at a variety of levels of morphological and syntactic structure.

The Results of the Experiments

- These results indicate that transitional probabilities can provide an important cue for identifying constituent structure from word sequences.
- The experiments provide data only on syntax learning by adults.
- However, when taken together with Saffran et al.'s research on infant identification of word boundaries, they strongly suggest that Bayesian inference of the kind employed in ML methods in NLP plays a significant role in human language acquisition at a variety of levels of morphological and syntactic structure.

Bayesian Bootstrapping in Language Acquisition

- This work also gives credence to a boot strapping view of language learning on which information theoretic methods yield an initial classification of linguistic entities that can then be used to construct successive levels of representation.
- Each previous cycle of learning provides a set of structural constraints on the entities out of which the next stage is developed by the same kinds of Bayesian inference.
- If this view is sustained by further research, then the weak bias model of language learning proposed in Lappin (2005), Lappin and Shieber (2007), and Clark (2004) will achieve psychological as well as computational credibility.

Bayesian Bootstrapping in Language Acquisition

- This work also gives credence to a boot strapping view of language learning on which information theoretic methods yield an initial classification of linguistic entities that can then be used to construct successive levels of representation.
- Each previous cycle of learning provides a set of structural constraints on the entities out of which the next stage is developed by the same kinds of Bayesian inference.
- If this view is sustained by further research, then the weak bias model of language learning proposed in Lappin (2005), Lappin and Shieber (2007), and Clark (2004) will achieve psychological as well as computational credibility.

Bayesian Bootstrapping in Language Acquisition

- This work also gives credence to a boot strapping view of language learning on which information theoretic methods yield an initial classification of linguistic entities that can then be used to construct successive levels of representation.
- Each previous cycle of learning provides a set of structural constraints on the entities out of which the next stage is developed by the same kinds of Bayesian inference.
- If this view is sustained by further research, then the weak bias model of language learning proposed in Lappin (2005), Lappin and Shieber (2007), and Clark (2004) will achieve psychological as well as computational credibility.

Conclusions

- The FOXP2 gene plays a central role in regulating the development of neural connections required for verbal, as well as non-verbal learning.
- It is not a "language gene", and the language disorder associated with point mutation of the gene in the KE family does not indicate the existence of a genetically coded language faculty.
- Important aspects of SLI are heritable and appear to have genetic causes.

Conclusions

- The FOXP2 gene plays a central role in regulating the development of neural connections required for verbal, as well as non-verbal learning.
- It is not a "language gene", and the language disorder associated with point mutation of the gene in the KE family does not indicate the existence of a genetically coded language faculty.
- Important aspects of SLI are heritable and appear to have genetic causes.

Conclusions

- The FOXP2 gene plays a central role in regulating the development of neural connections required for verbal, as well as non-verbal learning.
- It is not a "language gene", and the language disorder associated with point mutation of the gene in the KE family does not indicate the existence of a genetically coded language faculty.
- Important aspects of SLI are heritable and appear to have genetic causes.

Conclusions

- There is some evidence to suggest that these cause involve a working memory deficit, and (G-)SLI does not motivate a domain specific heritable language module.
- Recent psycholinguistic research reveals that Bayesian inference plays a significant role in identifying linguistic structure at a variety of levels.
- This work lends credibility to a bootstrapped view of language acquisition in which statistical learning produces successive levels of linguistic representation, each constraining the elements from which its successor is constructed.

Conclusions

- There is some evidence to suggest that these cause involve a working memory deficit, and (G-)SLI does not motivate a domain specific heritable language module.
- Recent psycholinguistic research reveals that Bayesian inference plays a significant role in identifying linguistic structure at a variety of levels.
- This work lends credibility to a bootstrapped view of language acquisition in which statistical learning produces successive levels of linguistic representation, each constraining the elements from which its successor is constructed.

Conclusions

- There is some evidence to suggest that these cause involve a working memory deficit, and (G-)SLI does not motivate a domain specific heritable language module.
- Recent psycholinguistic research reveals that Bayesian inference plays a significant role in identifying linguistic structure at a variety of levels.
- This work lends credibility to a bootstrapped view of language acquisition in which statistical learning produces successive levels of linguistic representation, each constraining the elements from which its successor is constructed.