

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE LETRAS

PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

ESTUDOS DA LINGUAGEM: AQUISIÇÃO DA LINGUAGEM

A EXPRESSÃO DO FOXP2

Uma introdução ao estudo das relações entre genes e linguagem

RENATA ROCHA FERNANDES GONÇALVES

Dissertação apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Mestre em Letras – Estudos da Linguagem.

Orientador: Prof^a. Dr^a. Luciene Juliano Simões

Porto Alegre

2006

AGRADECIMENTOS

À minha orientadora, Prof^ª. Dr^ª. Luciene Juliano Simões, pela compreensão, pelo estímulo, pela confiança e, principalmente, pelo apoio incondicional oferecido ao longo da minha trajetória na Universidade, sempre enriquecendo e deixando abertos os caminhos.

Aos professores da Graduação e do Programa de Pós-Graduação em Letras da UFRGS, pela convivência, pelo ensino e pelo conflito, que me fizeram crescer.

À CAPES, pela bolsa concedida e pelo Portal Periódicos, sem o qual este trabalho jamais poderia ter sido realizado.

Ao Christian, pela leitura atenta, pela palavra segura e pelo gesto carinhoso.

À minha mãe, de quem herdei o ímpeto de fazer, e ao meu pai, que me transmitiu o desejo de saber, pelo incentivo sempre renovado.

LISTA DE ABREVIATURAS

A – Adenina
ACC – Acusativo
A-P – Articulatório-perceptual
C – Citosina
C_{HL} – Sistema computacional humano
C-I – Conceitual-intencional
DAT – Dativo
DNA – Ácido desoxirribonucléico
DZ – Dizigótico
FF – Forma fonética
FL – Forma lógica
G – Guanina
GEN – Genitivo
GU – Gramática Universal
LSN – *Lenguaje de Signos Nicaragüense*
M1 – Área motora primária
MZ – Monozigótico
NOM – Nominativo
OBL – Oblíquo
PET – Tomografia por emissão de pósitrons
PM – Programa Minimalista
QI – Quociente de inteligência
RNA – Ácido ribonucléico
S_∅ – Estado inicial da faculdade da linguagem
S_e – Estado final da faculdade da linguagem
SLI – *Specific language impairment*
SNP – *Single nucleotide polymorphism*
SW – Síndrome de Williams
T – Timina
U – Uracila
WAIS – *Wechsler Adult Intelligence Scale*

LISTA DE FIGURAS, QUADROS E TABELAS

- Figura 1 – Arquitetura da faculdade da linguagem, segundo a concepção de Chomsky (1995)
- Figura 2 – Célula, núcleo celular e cromossomos
- Figura 3 – Cariótipo humano
- Figura 4 – Organização do DNA
- Figura 5 – Processo de *splicing*: remoção de íntrons na formação final do RNA
- Figura 6 – Código genético
- Figura 7 – Mecanismo de *splicing* alternativo
- Figura 8 – Processo de diferenciação celular
- Figura 9 – Mecanismo de ação dos fatores de transcrição
- Figura 10 – Modelo de limiar para heranças complexas
- Figura 11 – Heredograma da família KE
- Figura 12 – Identificação da região SPCH1
- Figura 13 – Resultados do teste de julgamento de gramaticalidade
- Figura 14 – Resultados do teste de flexão verbal
- Figura 15 – Estímulo para produção de narrativa curta
- Figura 16 – Habilidades verbais e visuo-espaciais na síndrome de Williams
- Figura 17 – Habilidades visuo-espaciais na síndrome de Williams e na síndrome de Down
- Figura 18 – Resultados do teste de praxia oral da família KE
- Figura 19 – Resultado dos testes de morfologia com palavras e pseudopalavras
- Figura 20 – Principais áreas do córtex cerebral associadas à linguagem
- Figura 21 – Distribuição somatotópica das áreas motoras no córtex cerebral
- Figura 22 – Distribuição do córtex e áreas subcorticais, destacando os núcleos da base
-
- Quadro 1 – Marcação de Caso no português e inglês
- Quadro 2 – Correlações entre desenvolvimento lingüístico e desenvolvimento cerebral
- Quadro 3 – Erros na flexão de plural de pseudopalavras
- Quadro 4 – Uso de pronomes, sintagmas e anáforas em teste de produção de narrativa curta
- Quadro 5 – Modelo de dupla dissociação da linguagem
-
- Tabela 1 – Frequência dos tipos de erros no teste de flexão verbal
- Tabela 2 – Produção e compreensão de marcas morfológicas flexionais e derivacionais
- Tabela 3 – QI verbal e não-verbal da família KE

RESUMO

A aquisição da linguagem é um processo invariavelmente dependente da exposição ao ambiente. Contudo, a mais famosa e controversa hipótese desenvolvida por Chomsky sustenta que a criança traz para o processo de aprendizagem da língua um conjunto de recursos específicos – a Gramática Universal – para superar a pobreza de estímulos do ambiente e construir uma gramática final adequada, composta por princípios universais e parâmetros fixados de acordo com a experiência na comunidade lingüística. A Gramática Universal é, assim, entendida como uma dotação genética, necessária, mas não suficiente, para a aquisição da linguagem. Para além da hipótese do inatismo – central à tradição gerativista – a relação entre genes e linguagem vem sendo investigada há muitos anos, com evidências significativas de transmissão hereditária e etiologia genética para diversos tipos de distúrbios de linguagem. Mais recentemente, em 2001, foi descoberto o primeiro caso de um gene, o *FOXP2*, implicado na habilidade de adquirir e desenvolver a fala e a linguagem. Este trabalho parte de uma introdução ao processo de aquisição da linguagem em uma perspectiva inatista (Capítulo I) para focar os aspectos centrais à organização biológica da faculdade da linguagem (Capítulo II), propondo, a partir da análise dos efeitos da mutação do gene *FOXP2*, uma discussão introdutória sobre como o estudo dos distúrbios do desenvolvimento da linguagem, avaliados a partir de suas características fenotípica e genotípica, pode contribuir para a exploração de questões fundamentais acerca do desenvolvimento lingüístico e cognitivo (Capítulo III).

ABSTRACT

Language acquisition invariably depends on environmental exposure. However, Chomsky's most famous and controversial hypothesis assumes that children bring to the process of language learning a set of specific resources – the Universal Grammar – in order to overcome difficulties imposed by poverty of stimulus and develop an adequate grammatical representation, composed of universal principles and parameters fixed by experience. Universal Grammar is, therefore, understood as a genetic make up, necessary but not sufficient for language acquisition. In addition to the innateness hypothesis – a central aspect of the generative tradition – the relationship between genes and language has been investigated for several years, with significant evidence of hereditary transmission and genetic etiology for various language disorders. More recently, in 2001, studies described the first case of a gene, *FOXP2*, thought to be implicated in our ability to acquire speech and language. This work begins with an introduction to the process of language acquisition according to the innateness hypothesis (Chapter I); this discussion is followed by a description of the main aspects concerning the biological organization of the language faculty (Chapter II). Finally, based on an analysis of the impact of the *FOXP2* mutation, we present an introductory debate on how developmental language disorders, evaluated based on their phenotype and genotype, can contribute to the advancement of fundamental questions concerning linguistic and cognitive development (Chapter III).

SUMÁRIO

I – A GRAMÁTICA UNIVERSAL NA ERA DA GENÔMICA

Fronteiras da Lingüística.....	9
Superando barreiras.....	21
Órgão da linguagem.....	29

II – DOS GENES À LINGUAGEM

Com quantos genes se faz um ser humano?	37
Genes e proteínas orquestrando o desenvolvimento.....	47
Moléculas da mente.....	53

III – ENFIM UM GENE PARA A LINGUAGEM. E DAÍ?

A história de uma descoberta.....	64
A definição do déficit nuclear como centro da discórdia.....	72
Por dentro do cérebro dos KE.....	104

CONCLUSÃO.....	119
----------------	-----

Human cognitive systems, when seriously investigated, prove to be no less marvelous and intricate than the physical structures that develop in the life of the organism. Why, then, should we not study the acquisition of a cognitive structure such as language more or less as we study some complex bodily organ?

(Noam Chomsky, 1975)

CAPÍTULO I

A GRAMÁTICA UNIVERSAL NA ERA DA GENÔMICA

Fronteiras da Lingüística

Tradicionalmente, o estudo da linguagem revela-se de interesse a diferentes áreas de investigação, que se expandem da filosofia à biologia. A própria história dos estudos lingüísticos nas últimas décadas parece confirmar em muito essa vocação: de uma lingüística sistêmica – voltada ao formal e ao abstrato e fechada em seu próprio objeto – emergiu uma lingüística aberta a intersecções com outras ciências, capaz de preservar seus métodos e objeto, mas também de revitalizar seus próprios questionamentos e reavaliar antigos postulados.

Essa virada deveu-se, em boa medida, ao desenvolvimento da lingüística gerativa, a partir do final da década de 1950. A hipótese de que o conhecimento humano acerca das línguas naturais é resultado de uma capacidade biologicamente determinada, específica à espécie, mantém-se ao longo dos anos como o aspecto mais controverso da teoria gerativa de Noam Chomsky e é, ainda hoje, o elemento que conecta mais fortemente a teoria lingüística à biologia e à psicologia (Jackendoff, 2002). Controvérsias à parte, a teoria gerativa ofereceu à lingüística um novo modelo de abordagem e investigação dos fenômenos da língua, ampliando o espectro e o volume de pesquisas realizadas na área.

O entrelaçamento com outras áreas científicas esteve na base do estabelecimento da lingüística como disciplina acadêmica, nos Estados Unidos, onde inicialmente desenvolveu-se ancorada em relações com a antropologia, através dos trabalhos de Franz Boas (1858-1942) e Edward Sapir (1884-1939), e com a psicologia, a partir da obra de Leonard Bloomfield (1887-

1949). Com o posterior desenvolvimento da corrente gerativista, a lingüística voltaria a dialogar com outras ciências, principalmente a partir do momento em que o inatismo (discutido adiante) colocou-se como paradigma central à teoria, erguendo assim as ciências biológicas à posição de potencial interlocutor.

A inserção do estudo da linguagem como parte fundamental das investigações acerca da cognição humana – em oposição ao estudo da linguagem enquanto processo histórico ou objeto social – deve-se originalmente a um grupo de pesquisadores que, no que tange ao modo de investigação dos fenômenos lingüísticos, sustentou posições diametralmente opostas àquelas que são hoje advogadas pela maioria dos interessados nas relações entre cérebro e linguagem. Partidários de uma abordagem ambientalista e antimentalista, Bloomfield e seus seguidores desenvolveram suas obras dentro do espírito behaviorista, sustentando a hipótese de que o comportamento humano – incluindo aí a linguagem – seria condicionado a partir de reforços positivos a determinadas ações executadas em resposta a estímulos do ambiente. Na tentativa de explicar o comportamento através da formação de hábitos, os behavioristas tributavam os “hábitos verbais” a um simples processo de imitação por parte da criança.

Foi justamente em oposição a essa proposta que Chomsky tornou inicialmente conhecidas suas idéias acerca da natureza da linguagem humana. Em resenha negativa sobre o livro *Verbal Behavior* (1959), de B. F. Skinner, Chomsky, além de combater a extrapolação de um modelo animal para o estudo de uma faculdade cognitiva humana, argumentava que o chamado comportamento verbal não poderia ser encarado apenas como o produto de um conjunto de fatores ambientais. Segundo Chomsky, para poder explicar o conhecimento lingüístico alcançado por um falante adulto de uma língua natural qualquer, é preciso fazer referência a construtos mentais previamente disponíveis ao homem, diretamente envolvidos no uso e na compreensão da linguagem. A idéia de que o conhecimento lingüístico resulta de uma capacidade inata específica – e não de um comportamento indutivo a partir da mera

observação da linguagem – viria a se tornar um dos pontos centrais da tradição gerativista desenvolvida a partir de então.

O programa gerativista compõe-se de duas frentes de trabalho, que se complementam no objetivo final de compreender a natureza da linguagem humana. Como comenta Raposo (1992), o objeto do empreendimento gerativo é, por um lado, a caracterização das gramáticas particulares dos indivíduos (i.e., as várias línguas humanas) e, por outro lado, a caracterização da Gramática Universal (GU), entendida como um conjunto de propriedades inatas, biologicamente determinadas, de natureza especificamente lingüística (ou seja, não partilhada por nenhum outro sistema cognitivo), e cujo desenvolvimento e maturação, em interação com o meio ambiente, determina uma gramática particular na mente do adulto.

A publicação de *Syntactic Structures*, em 1957, é considerada um divisor de águas na história da lingüística (Weedwood, 2002), que a partir daí passa a se distanciar das posições estruturalistas e behavioristas, estabelecendo uma vertente nos estudos da linguagem em maior sintonia com o desenvolvimento das ciências cognitivas. Conforme observa Jackendoff (1997), a posição básica da lingüística gerativa é a de que a linguagem não deve ser estudada como algum tipo de fenômeno abstrato ou artefato social, já que sua natureza está, na verdade, no modo como ela é empregada e compreendida rotineiramente por seres humanos.

Um dos principais argumentos sustentados por Chomsky à época para justificar a necessidade de um construto teórico como a GU girava em torno da chamada pobreza de estímulos: o ambiente seria incapaz de fornecer à criança dados suficientes para que ela derivasse, exclusivamente com base nesses dados, toda a gramática de sua língua. Na ausência de um contexto patológico, ao final do processo de aquisição, por volta dos quatro anos de idade, o conhecimento de que toda e qualquer criança dispõe acerca de sua própria língua ultrapassa largamente os dados observáveis no ambiente. Assim, o estímulo do ambiente é dito “pobre” em pelo menos três sentidos:

- (1) A experiência da criança com a língua é finita, ou seja, ela não é exposta a todos os dados da língua, mas suas habilidades lingüísticas ao final do processo superam em muito esse *input* inicial (i.e., a criança é capaz de produzir e compreender frases que jamais ouviu anteriormente).
- (2) A experiência da criança com a língua não consiste em um conjunto hermeticamente organizado de dados, disponibilizados em ordem crescente de complexidade e dirigidos especificamente a ela. Antes, essa experiência com a língua resulta, ao menos em parte, do entreouvir de dados fragmentados, frases incompletas e referências perdidas.
- (3) Os dados por si só, especialmente em função da ausência de *feedback* negativo (ou seja, da ausência de informações sobre aquilo que não ocorre ou não é permitido na gramática da língua), não oferecem o substrato necessário para a indução das diversas regras e generalizações manifestas no estágio final da aquisição.

Características próprias à organização e ao funcionamento da linguagem somam-se aos argumentos em torno da pobreza de estímulos na construção da hipótese chomskyana em favor da existência de mecanismos inatos subjacentes ao processo de aquisição da linguagem. A infinitude discreta das línguas naturais seria uma característica importante nesse sentido, na medida em que todas as línguas humanas se constituem como um conjunto limitado de elementos capaz de – combinatoriamente – gerar uma quantidade ilimitada de expressões. Segundo Chomsky, apenas um conhecimento inato, geneticamente determinado, pode explicar o fato de que a criança pressupõe certas propriedades combinatórias da linguagem sem receber qualquer instrução explícita.

A linguagem humana está baseada numa propriedade elementar que parece também ela ser biologicamente isolada: a propriedade da infinitude discreta, que em sua forma mais pura é exibida pelos números naturais 1, 2, 3, ... As crianças não aprendem esta propriedade; se a mente não possuísse já de antemão os princípios

básicos, não haveria quantidade de evidência capaz de provê-los. Do mesmo modo, nenhuma criança precisa aprender que existem frases com três palavras e frases com quatro palavras, mas não com três palavras e meia, e que esse número pode ir aumentando sem ter fim; é sempre possível construir uma frase mais complexa, com uma forma e um sentido definidos. Tal conhecimento nos chega necessariamente “originário da mão da natureza”, nas palavras de David Hume, como parte da nossa dotação biológica (Chomsky, 1997).

Assim, Chomsky enfatiza que o conhecimento humano sobre a própria língua estende-se muito além de tudo aquilo que é ou pode ser observado pelo indivíduo ao longo da vida. Com base em uma experiência limitada, cada falante constrói um sistema produtivo – a gramática de sua língua – que não apenas incorpora as instâncias concretas de vivência lingüística, mas que também permite a produção e a compreensão de um número ilimitado de novos enunciados.

Além da pobreza de estímulos e da infinitude discreta, o gerativismo ressalta a similitude entre as diferentes línguas naturais como indicativo de que essas seriam produto de uma faculdade humana comum à espécie, e não meros artefatos culturais. As diversas línguas em todo o mundo compartilham uma estrutura lógica comum subjacente às diferenças que emergem na superfície, apresentando, por exemplo, as categorias de nome, verbo e tempo; um elemento formador de perguntas, como *wh-* (ou *qu-*); a existência de pronomes; palavras para designar cores; e assim por diante. Um dos universais lingüísticos mais pesquisados diz respeito à ordem das palavras nas frases. A ordem básica das frases em português é SVO, ou seja, sujeito-verbo-objeto (pode haver alteração, mas esse é o padrão básico do português, do inglês, do francês e da maioria das línguas ocidentais). Na sua imensa maioria, as línguas seguem padrões nos quais o sujeito antecede o objeto (SVO, VSO, SOV), sendo raros os casos de línguas nas quais o objeto antecede o sujeito (VOS, OVS, OSV). Assim, os universais lingüísticos não estabelecem um padrão único, mas sim os padrões possíveis.

Considere a categoria gramatical Caso. O conceito tradicional de Caso incorpora o Caso semântico (atribuição de papéis a cada um dos argumentos) e o Caso morfológico (paradigma de morfemas que determinam as funções sintáticas dos elementos da oração).

- Caso semântico

[O veterinário_{agente}] operou [o porco_{tema}] [com uma faca_{instrumento}] [no chiqueiro_{locativo}].

- Caso morfológico

Puer_{NOM} puell**am**_{ACC} amat.

Puella_{NOM} puer**am**_{ACC} amat.

Os morfemas de Caso (em negrito no exemplo) marcam as funções gramaticais dos sintagmas da frase (sujeito e objeto, neste caso) e, através deles, reconhece-se o papel temático dos elementos (quem ama e quem é amado). Embora todas as línguas apresentem a noção de Caso, nem todas deixam a marca visível, como o Latim, que tem morfemas específicos para marcar caso nominativo, acusativo, dativo e ablativo. Em português, e em muitas outras línguas, não há marca morfológica de Caso.

Todas as línguas são idênticas por disporem da categoria gramatical Caso. O que pode ser diferente de uma língua para outra é a forma como esta categoria se expressa: em algumas, concretamente, por meio de um paradigma mais ou menos rico de morfemas; em outras, abstratamente. Daí a noção de Caso abstrato (Miotto et al., 2000:114).

Assim, Caso é um fenômeno universal, comum a todas as línguas e de natureza essencialmente sintática. Em português, a existência de Caso é comprovada através de sua manifestação morfológica (residual) no sistema pronominal. Em inglês, há marcação de Caso genitivo e também resíduos no sistema pronominal (Quadro 1):

Português	Inglês
Eu _{NOM} ofereci- <i>lhe</i> _{DAT} o livro.	John's _{GEN} girlfriend is called Mary.
Ele _{NOM} comprou- <i>o</i> _{ACC} para <i>mim</i> _{OBL} .	She _{NOM} disliked him _{ACC} .

Quadro 1. Exemplos da marcação de Caso no português e no inglês.

Ainda em termos de universalidade, o desenvolvimento do *pidgin* em crioulo na sucessão de gerações e as sofisticadas gramáticas desenvolvidas pelos usuários das línguas de sinais também corroboram a idéia de que todas as línguas compartilham um arcabouço cognitivo comum (Anderson & Lightfoot, 1999).

O *pidgin* é uma língua de contato, que surge quando povos falantes de línguas mutuamente ininteligíveis entram em contato estreito, ou seja, quando têm necessidade de se comunicar uns com os outros, como ocorreu durante a exploração do novo mundo pelos europeus. Naturalmente, a condição inicial dessa comunicação é bastante precária, e a língua que surge a partir daí – o *pidgin* – tem um léxico reduzido e uma morfologia quase inexistente, permitindo apenas uma comunicação mínima. Contudo, se esse grupo misto de colonizadores e colonizados começar a se consolidar, dando origem a uma primeira geração de nativos da nova comunidade, o *pidgin* passa a ser língua materna dessas crianças: tem-se, então, o crioulo. Logo, a língua crioula é um *pidgin* que foi adquirido como língua nativa; é um *ex-pidgin*, que virou língua materna de uma comunidade. Como consequência, as crianças incorporam e desenvolvem o *pidgin* dos pais, observando-se um processo de gramaticalização da língua, ampliação do vocabulário e enriquecimento das possibilidades comunicativas (Couto, 1996).

Esse fenômeno da transformação de uma língua provisória e simplificada numa língua bem estabelecida e complexa pode ser visto também em comunidades de surdos. Ao contrário do que diz o senso comum, as línguas de sinais não são pantomimas, gestos memorizados ou criados por professores; elas são encontradas em todas as comunidades de deficientes auditivos e compartilham com a linguagem falada uma série de complexidades gramaticais, expressas através de sutilezas gestuais frequentemente imperceptíveis ao olho não-treinado. Como exemplifica Pinker (2002), há alguns anos não havia nenhuma língua de sinais na Nicarágua, porque seus deficientes auditivos eram mantidos isolados uns dos outros. Após

uma reforma do sistema de ensino, na década de 1980, foram criadas as primeiras escolas para surdos. Apesar do enfoque educativo ser na leitura labial e na fala, as crianças naturalmente desenvolveram seu próprio sistema de sinais, acumulando gestos que se consolidaram como a *lenguaje de signos nicaragüense* (LSN).

Como essa língua foi desenvolvida por jovens que, à época, já tinham em torno de 10 anos de idade, essa língua não se desenvolveu plenamente: é um *pidgin* e permanece sendo usada, com variados graus de fluência, pela comunidade surda nicaragüense que tem hoje cerca de 30 anos. Contudo, crianças que entraram na escola, aos quatro ou cinco anos de idade, quando a LSN já vigorava, foram expostas ao *pidgin* das mais velhas e introduziram estratégias gramaticais antes ausentes, dando origem a um crioulo – o *idioma de signos nicaragüense* – muito mais expressivo e uniformemente empregado pela comunidade surda a partir de então.

No que tange ao problema específico da aquisição da linguagem, a abordagem chomskyana deriva da própria concepção mentalista e composicional que subjaz ao programa gerativista¹. Desde o princípio de seu desenvolvimento, a teoria gerativa propôs-se a construir um modelo de organização e funcionamento da linguagem humana que fosse além de uma adequação observacional e descritiva e atingisse uma adequação explanatória. Tal adequação seria alcançada justamente quando a teoria fosse capaz de dar conta do fato de que os princípios postulados para a gramática internalizada dos falantes adultos devessem ser adquiridos, de algum modo, pelas crianças (Haegeman, 1991). Ou seja, para atingir a adequação explanatória seria preciso caracterizar o estado inicial (S_{\emptyset}) da faculdade da linguagem e, então, mapear a passagem para o estado final (S_e). Tanto princípios gerais

¹ É importante observar que o fenômeno da aquisição da linguagem não constitui um objeto de investigação próprio ao programa gerativista. No entanto, o objetivo de alcançar uma adequação explanatória fez da imagem da criança aprendendo sua língua um tema central e recorrente na tradição gerativista, “de maneira que se torna difícil discriminar sua teoria da linguagem de sua visão da aquisição da linguagem” (Scarpa, 2000:207).

(universais) lingüísticos quanto a experiência particular (local) com a língua devem constituir a base de investigação desse mapeamento ($S_{\emptyset} \rightarrow S_e$).

Em *Aspects of the Theory of Syntax* (1965), Chomsky afirmava que a aquisição da linguagem por parte da criança envolveria essencialmente o desenvolvimento de um sistema internalizado de regras sobre a formação, o uso e a compreensão de frases em uma dada língua. A construção dessas regras deve ser entendida como um processo não-intencional, isto é, a criança não participa do processo aprendendo fatos sobre a própria língua e derivando regras a partir daí. A gramática internalizada é inconsciente e aprender uma língua corresponderia, nesse sentido, a “vir a saber que” (Jackendoff, 2002). Ao postular que a criança possui um sistema internalizado de regras, no entanto, pouco ou nada se avança em termos explanatórios, já que resta saber como esse sistema foi alcançado.

Os modelos teóricos anteriores à Teoria de Princípios & Parâmetros (1981) – por estarem voltados à definição de regras específicas que respondessem pelas estruturas encontradas em línguas particulares – pouco contribuíram para esclarecer essa questão, deixando a adequação explanatória como uma meta ainda por ser atingida. Como ressalta Meisel (1997:22), “quanto mais elaboradas se tornaram as regras – transformações e outras – menos óbvio se tornou o modo pelo qual as crianças poderiam adquirir esses sistemas de regras”.

Chomsky (1995) afirma que o modelo proposto com a Teoria de Princípios & Parâmetros constituiu um rompimento radical, não apenas com a tradição gerativista, mas com toda a história de estudos da linguagem, na medida em que sustentou que uma idéia central à tradição (incorporada sem maiores alterações pela lingüística gerativa), qual seja, a de que a linguagem é constituída por um conjunto de regras, estava fundamentalmente errada. Pela nova abordagem proposta, que norteia as discussões deste trabalho, as línguas não se

organizam em torno de regras específicas, mas em função de princípios universais e de um arranjo finito de opções – os parâmetros – em torno dos quais esses princípios podem ser organizados. Os princípios são, em geral, restrições (*constraints*), proposições negativas sobre a organização da linguagem (aquilo que não é permitido). Os parâmetros, por sua vez, são propriedades que variam de língua para língua, constituindo-se de opções limitadas para o arranjo de categorias funcionais.

Para o campo da aquisição, essa abordagem resultou na abertura de novas possibilidades investigativas. Em especial, os investigadores passaram a buscar a caracterização do estado inicial (S_0) – a Gramática Universal – e a descrição dos parâmetros sujeitos à fixação em função da experiência (i.e., do contato com a língua do ambiente). Em conjunto, essas “duas gramáticas” resultam no estado final do conhecimento lingüístico do indivíduo. “*We call the theory of the state attained its grammar and the theory of the initial state Universal Grammar*” (Chomsky, 1995:14). Assim, com o modelo de princípios e parâmetros, o foco da pesquisa em aquisição passou a ser a descrição dos elementos dos quais a criança precisaria ter conhecimento, de antemão, para que, a partir dos dados do *input*, fosse capaz de chegar ao estado final (S_e), compartilhado por sua comunidade lingüística.

Subseqüentemente, Chomsky desenvolve o Programa Minimalista (PM) (Chomsky, 1995) que propõe uma eliminação de excedentes teóricos encontrados em modelos anteriores, como a Teoria Padrão (1965) e a Teoria de Princípios e Parâmetros (1981). Embora assuma a abordagem central deste último, o PM introduz princípios de economia que terminam por reestruturar o modelo de gramática e a própria caracterização da faculdade da linguagem. Não se trata de um novo paradigma teórico, mas de um conjunto de orientações para evitar a postulação de entidades teóricas que não sejam conceitualmente necessárias à lógica da teoria.

O PM mantém, assim, algumas das posições sustentadas por versões anteriores da teoria chomskyana. Entre elas, a de que existe um componente no cérebro humano dedicado exclusivamente à faculdade da linguagem, interagindo com outros sistemas. A faculdade da linguagem seria constituída por dois componentes: um sistema cognitivo (foco principal de interesse do PM), que armazena informações, e um sistema de performance, que acessa e utiliza essas informações de formas variadas. Mantém-se a idéia milenar de que a linguagem estaria organizada em torno de relações de som e significado; logo, o sistema cognitivo e o sistema de performance devem interagir um com o outro através de dois níveis de interface: um sistema articulatório-perceptual A-P e um sistema conceitual-intencional C-I.

Um língua particular-L, por sua vez, organiza-se em torno de um sistema computacional (C_{HL}), comum às línguas humanas, e de um léxico, variável nas diferentes comunidades lingüísticas. O léxico especifica o conjunto de elementos que o C_{HL} pode selecionar e integrar para formar expressões lingüísticas, as quais são compostas por pares de som e significado (π, λ) , que são convertidos em uma representação abstrata, a forma fonética (FF) e a forma lógica (FL), respectivamente, as quais são interpretadas nas devidas interfaces, A-P e C-I (Fig. 1).

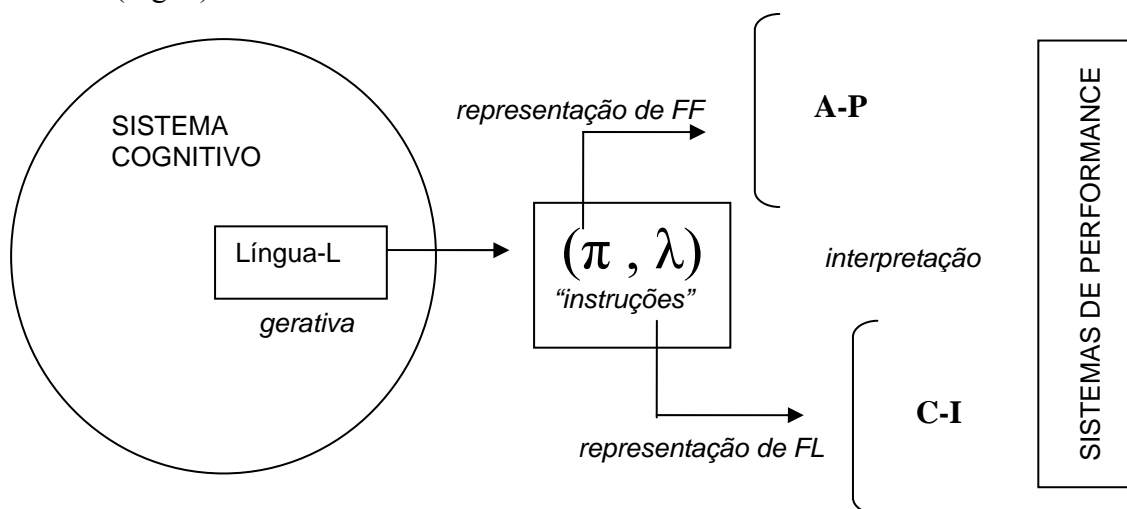


Figura 1. Uma língua particular (Língua-L) representa a organização final da Gramática Universal, após a fixação de parâmetros. A Língua-L é um sistema gerativo que produz pares (π, λ) que são interpretados nas interfaces articulatório-perceptual (A-P) e conceitual-intencional (C-I), respectivamente, como instruções para os sistemas de performance. π é uma representação da forma fonológica (FF) e λ é uma representação da forma lógica (FL) de uma determinada expressão da língua.

A especificidade da linguagem, enquanto parte do sistema cognitivo, advém da natureza dos procedimentos computacionais de C_{HL} (universais), dos traços morfológico-formais do léxico (variáveis) e das *bare output conditions* – condições de legibilidade impostas por sistemas externos às interfaces, que farão uso da informação fornecida por C_{HL} . Assim, p.ex., as derivações (computações) executadas pela Língua-L devem ser passíveis de externalização pelo aparato motor humano. A capacidade respiratória impõe limites sobre a extensão possível das frases faladas, assim como a memória de trabalho restringe a possibilidade de compreensão até um certo nível de complexidade (Hauser et al., 2002). “Os sistemas dentro dos quais a faculdade da linguagem se encaixa têm de ser capazes de ‘ler’ as expressões da língua e usá-las como ‘instruções’ para o pensamento e a ação” (Chomsky, 1998:13).

A faculdade da linguagem inclui, assim, um componente computacional (C_{HL}) interno, que interage com pelo menos dois subsistemas (A-P e C-I), os quais, por sua vez, interagem com sistemas externos, os quais lhe impõem restrições. C_{HL} é um componente-chave (sintático) da faculdade da linguagem que gera representações internas e as mapeia para a interface A-P, através de uma representação fonológica, e para C-I, através de uma representação semântica. As propriedades computacionais específicas de C_{HL} são discutíveis, mas uma das propriedades consideradas centrais é a recursividade, a capacidade de tomar um conjunto restrito de elementos e gerar um arranjo potencialmente infinito de expressões. Além de C_{HL} , a faculdade da linguagem é constituída também por um léxico, que especifica os itens que entram no sistema computacional, sendo entendido como uma lista de exceções e idiosincrasias, particular a cada uma das línguas, contendo tudo aquilo que não pode ser derivado de princípios gerais da GU, incluindo aspectos fonológicos e morfológicos e opções de valores paramétricos.

Superando barreiras

O fato de que, de alguma forma, as crianças chegam, de modo razoavelmente uniforme, ao final do processo de aquisição com a interpretação correta para os dados fragmentários disponíveis no ambiente reforça, segundo a teoria, a idéia da existência de um mecanismo inato capaz de elaborar corretamente os dados lingüísticos primários, gerando uma gramática específica. Lightfoot (1989) propôs o seguinte modelo, que estabelece um interessante paralelo entre as terminologias lingüística e genética:

- a. *trigger* (genótipo → fenótipo);
- b. dados lingüísticos primários (Gramática Universal → gramática).

De acordo com essa proposta, a Gramática Universal (GU) seria a estrutura presente na mente/cérebro da criança que conteria os aspectos do genótipo presumivelmente relevantes para o desenvolvimento da linguagem, enquanto a gramática particular da língua seria considerada uma propriedade fenotípica, parte do arranjo mental que caracteriza o estado maduro da capacidade lingüística de um falante do português, do inglês ou de outra língua qualquer. Os dados lingüísticos primários são aqueles dados aos quais as crianças são expostas e que realmente determinam (*trigger*) algum aspecto de suas gramáticas, tendo algum efeito em longo prazo. Assim, lingüisticamente, todos os seres humanos compartilhariam um mesmo arcabouço de adaptações funcionais para a linguagem – podendo tornar-se falantes do húngaro, do alemão ou do mandarim – dependendo inteiramente do fator ambiental para tal determinação.

Segundo Bates (1999), os pesquisadores, hoje em dia, estão em comum acordo quanto ao fato de que genes e ambiente estariam interagindo na determinação de estruturas cognitivas complexas, como a linguagem. No entanto, a controvérsia *nature* versus *nurture*, ou seja, o debate em torno do papel da natureza e do ambiente, ou do biológico e do social, na

determinação desses comportamentos permanece no centro das discussões sobre o processo de aquisição da linguagem.

Dentro do paradigma gerativista, a posição de relevo dada ao ambiente no âmbito do behaviorismo foi drasticamente reduzida: de base e fundamento da aquisição, o ambiente lingüístico passou a desempenhar o papel de gatilho (*trigger*) necessário ao desencadeamento do processo. Dessa forma, o ambiente atuaria selecionando valores particulares para os parâmetros da GU; dito de outra forma, a GU seria a base para a aquisição de qualquer gramática humana, dada uma experiência lingüística adequada. Deve-se enfatizar que a GU não deve ser confundida com a gramática de qualquer língua em particular; ela é, sim, um conjunto de princípios através dos quais a criança é capaz de construir – com base nos dados limitados disponíveis no ambiente – o conhecimento gramatical pleno de sua própria língua.

Nessa perspectiva, dada a existência de um estágio inicial não-nulo, determinado pela GU e pelos parâmetros fixáveis, a criança não embarca no processo de aquisição como *tabula rasa*: para toda e qualquer língua, a faculdade da linguagem – entendida como um dos componentes do sistema cognitivo – coloca à sua disposição um conjunto X de princípios universais, sobre os quais deverão atuar um conjunto Y de parâmetros variáveis.

A aquisição é algo que acontece com a criança e não algo que a criança faz. E, embora o meio ambiente importe claramente, o curso geral do desenvolvimento e os traços básicos do que emerge são pré-determinados pelo estado inicial. Mas o estado inicial é uma posse comum aos homens. Tem de ser então que, em suas propriedades essenciais, as línguas são moldadas na mesma forma. (Chomsky, 1998:23)

Ao vincular a linguagem a mecanismos inatos, específicos à espécie humana, Chomsky situa seu objeto dentro de um domínio cognitivo e biológico. Admite que o ser humano, ao nascer, é dotado de um tipo de instrumental – a GU –, constituído por princípios universais da faculdade da linguagem necessários à passagem entre $S_{\emptyset} \rightarrow S_e$ e por um

conjunto de parâmetros a serem fixados pela experiência. Em suma, em relação à linguagem, é verdadeiro tanto afirmar que se trata de algo inato quanto de algo aprendido, sem que isso implique qualquer tipo de paradoxo: como diversos processos de desenvolvimento, a aquisição da linguagem envolve uma forma especial de aprendizagem, cujo curso é fortemente restrito por condições biológicas preestabelecidas (Jackendoff, 2002).

Sabe-se que, na ausência de um contexto patológico, todas as crianças aprendem com sucesso sua(s) língua(s) nativa(s), sem que recebam qualquer tipo de cuidado ou atenção especial para que alcancem esse objetivo. No entanto, até o estabelecimento definitivo de uma gramática internalizada, a fala infantil apresenta uma série de expressões desviantes e fragmentárias em relação à fala adulta, que de modo algum poderiam ser atribuídas à imitação da fala de seus pais. Décadas de estudo sobre o desenvolvimento da linguagem nas crianças mostraram não apenas que o processo de aquisição segue uma seqüência de estágios semelhante nas diferentes línguas e entre diferentes crianças, mas revelaram também que a ocorrência de erros na fala infantil não se dá aleatoriamente, de modo casual, mas antes demonstra um padrão de erros sistêmicos e recorrentes. Os três exemplos abaixo (extraídos do Childes, cf. MacWhinney, 1991) ilustram erros típicos do inglês infantil:

- a. *Eve sit floor.* (Eve, 1;7);
- b. *Where penny go?* (Adam, 2;4);
- c. *That truck fall down.* (Nina, 2;0).

Como enfatizado anteriormente, a dotação genética de uma GU é condição necessária, mas não suficiente ao desenvolvimento integral da linguagem, que irá depender da interação desta com os dados extraídos do ambiente. Segundo Jackendoff (2002), exemplos como esses indicam que as crianças possuem a capacidade de “inventar” uma gramática gerativa, com suas próprias regras de formação e interpretação das frases a partir dos dados lingüísticos

primários, mesmo quando essas regras ainda se mostram muito deficientes em seu modo de operar sobre os dados captados no ambiente.

No entanto, para atingir o estado final de desenvolvimento lingüístico, a criança precisa reinventar ou de algum modo redefinir os elementos dessa gramática inicial, de caráter provisório. Resta, então, a pergunta: se uma criança não sabe P na idade n , como essa criança vem a saber P na idade de $n + m$? A literatura aquisicionista gerativista apresenta pelo menos três encaminhamentos possíveis para responder a essa pergunta, os quais serão brevemente apresentados a seguir. Para fins de clareza expositiva, essas três alternativas serão aqui chamadas de hipótese da aprendizagem, da maturação cognitiva e da maturação lingüística.

A hipótese da aprendizagem foi desenvolvida por Nina Hyams (1986), a partir da questão da fixação de parâmetros – em especial, do parâmetro *pro-drop*. As línguas [+ *pro-drop*], como o italiano e o português brasileiro, são aquelas que aceitam um argumento-nulo, ou seja, aquelas nas quais não há obrigatoriedade dos sintagmas nominais de sujeito ou objeto serem preenchidos lexicalmente. Ao contrário, nas línguas [– *pro-drop*], como o inglês e o alemão, essas posições devem estar sempre ocupadas por um item lexical. Observando que, durante certo estágio do desenvolvimento lingüístico, as crianças falantes do inglês omitem a expressão do sujeito pronominal, obrigatório nessa língua, Hyams desenvolveu a seguinte proposta:

- a. Default [+ *pro-drop*]: a GU instrui as crianças a adotar a hipótese inicial de que sua língua deve ser do tipo do italiano ou português, ou seja, uma língua que admite sujeito nulo;
- b. Refixação [– *pro-drop*]: as crianças que estão adquirindo línguas como o inglês ou o alemão, que são línguas de sujeito obrigatório, têm de aprender, i.e., descobrir através do *input*, que sua língua não é de sujeito nulo, mas de sujeito obrigatório.

Assim, Hyams propõe uma hipótese baseada na capacidade de aprendizagem da criança, assumindo que esta dispõe de uma hipótese inicial (o parâmetro marcado positivamente) e precisa aprender – com base nos dados de sua língua – se essa hipótese inicial é adequada ou não para a gramática particular da língua que está adquirindo.

Dentre as críticas à hipótese da aprendizagem, a principal concentra-se em torno de uma lacuna explicativa – o chamado *triggering problem* – que pode ser assim descrito: admitindo-se que a frequência e a composição do *input* são constantes, por que, em muitos casos, os dados do *input* não são capazes de desencadear a aprendizagem de uma dada construção em um momento m , mas são capazes de fazê-lo em um momento $m+1$? Embora não haja uma resposta definitiva ao problema, uma solução possível aventada pelos defensores da hipótese refere-se à existência de um limiar de frequência (*frequency threshold*), ou seja, um número X de ocorrências relevantes de uma determinada construção deveria ser atingido para que a criança passasse a atentar a essa construção, eventualmente refixando o respectivo parâmetro.

As hipóteses maturacionistas (maturação cognitiva e maturação lingüística) propõem soluções alternativas a esse problema. A hipótese da maturação cognitiva atribui o conhecimento tardio (*late knowledge*) de certas construções gramaticais, como as frases sem sujeito no estágio inicial do inglês, a limitações impostas por sistemas cognitivos externos à faculdade da linguagem, que afetam o processamento dos dados lingüísticos. Assim, uma alternativa para solucionar o *triggering problem* consiste em invocar limitações cognitivas da criança.

Desenvolvida principalmente por Paul Bloom (1990), essa teoria afirma que as crianças de cerca de dois anos de idade, ao entrar no processo de aquisição, são dotadas de habilidades cognitivas e gramaticais iguais às dos adultos. Isso significa que todos os princípios gramaticais disponíveis na gramática de um falante adulto já se fazem presentes na

gramática do aprendiz, sendo as propriedades sintáticas particulares da língua rapidamente adquiridas com base no *input*. Assim, usando o mesmo exemplo, dado que crianças de dois anos adquirindo o inglês são expostas a uma língua [*- pro-drop*], isso implica – pelos pressupostos da teoria – que elas sabem que sujeitos, nessa gramática, são obrigatórios. Conseqüentemente, frases como “*want water*” devem ser o resultado de limitações de desempenho ou performance – não dependendo do conhecimento ou da competência lingüística do falante.

As limitações na produção verbal infantil, que levam à omissão de palavras e sintagmas na fala espontânea, sugeririam, assim, que as crianças são vulneráveis a limitações de processamento cognitivo, em especial a restrições quanto à capacidade da memória de trabalho. Como argumenta Bloom (1993), é fácil subestimar a competência do falante a partir de julgamentos de seu desempenho; o que faz a hipótese da maturação cognitiva, então, é pressupor a competência total e atribuir à imaturidade do cérebro em desenvolvimento as limitações da fala infantil.

Especificamente, Bloom sugere que a criança não seria capaz de processar ou produzir orações com um número muito grande de palavras. Por isso, deixaria de fora aqueles elementos que podem ser recuperados a partir do contexto. Isso explicaria a assimetria entre a omissão de sujeitos e objetos verificada em diversas línguas: as crianças omitiriam sujeitos mais freqüentemente porque o sujeito de uma frase é, geralmente, o tópico, o assunto, a informação já presente no contexto.

Com o crescimento, à medida que a capacidade de processamento se desenvolve e a de armazenamento presumivelmente se expande, a criança se tornaria capaz de produzir e compreender enunciados mais complexos e poderia, com isso, respeitar os princípios gramaticais de sua língua – como a obrigatoriedade de sujeitos.

Trata-se essencialmente de uma explicação não-lingüística, de boa aceitação no domínio da psicologia cognitiva, pela qual crianças e adultos compartilham as mesmas habilidades cognitivas, muito embora as crianças sofram restrições adicionais impostas pela condição ainda imatura do cérebro. Críticos dessa hipótese questionam a validade de seu princípio central – de que os efeitos da imaturidade de um domínio cognitivo particular (memória) possam afetar outros domínios (linguagem) – sugerindo que essa hipótese, para ser validada, deveria ser demonstrada em diferentes aspectos da cognição infantil.

A terceira hipótese para responder às perguntas anteriormente lançadas – se uma criança não sabe P na idade n , como essa criança vem a saber P na idade de $n+m$? e por que os dados do input não são suficientes em m , mas em $m+1$? – é a hipótese da maturação lingüística, defendida especialmente por Kenneth Wexler (1990), a partir do ideário original lançado por Eric Lenneberg ainda na década de 1960. Segundo essa hipótese, o processo de aquisição está na dependência direta de um processo de maturação da própria faculdade da linguagem, independentemente de outros domínios cognitivos: os estágios no desenvolvimento do conhecimento lingüístico seriam determinados pelo processo de maturação do cérebro, que afetaria propriedades específicas da GU. Logo, as operações da gramática – e, portanto, o desenvolvimento da linguagem no indivíduo – deveriam refletir o processo de desenvolvimento e constituição orgânicos do cérebro, isto é, o processo de maturação biológica deste órgão.

A principal evidência em favor dessa hipótese vem da existência de um período crítico para aquisição da linguagem, ou seja, de um período definido (e único) de tempo dentro do qual seria possível à criança adquirir normalmente a sua língua materna. A ausência de estímulo lingüístico relevante durante esse período – como observado nos raros casos relatados das chamadas crianças selvagens ou daquelas que, por diferentes motivos, foram

privadas do contato humano na infância – levaria a um desenvolvimento precário e irregular da linguagem.

“Por que as crianças começam normalmente a falar entre os 18 e os 28 meses?” – perguntava Lenneberg, em *Fundamentos Biológicos da Linguagem* (1967:153). Será que a criança precisa mudar aos 18 meses porque o seu ambiente costuma mudar nessa época, ou será que ela própria sofre mudanças importantes nesse estágio? Será o caso que a sociedade e os pais passam a se comportar de forma diferente quando a criança atinge 18 meses? – questiona o autor. “É bastante evidente que as diferenças mais importantes entre as fases pré-lingüística e pós-lingüística do desenvolvimento têm sua origem no indivíduo em crescimento e não no mundo externo ou nas mudanças na disponibilidade de estímulos” – conclui Lenneberg.

O início do período crítico seria delimitado por um estado de imaturidade cerebral e seu final marcado pela redução acentuada da plasticidade neuronal, simultânea ao processo de lateralização hemisférica da linguagem e à perda acentuada da capacidade de reorganização das funções cognitivas (em caso de lesão cerebral). Entre esses dois extremos, Lenneberg aponta ainda alguns estágios de correlação entre desenvolvimento lingüístico e cerebral, conforme apresentado no Quadro 2.

Desenvolvimento da linguagem	Desenvolvimento cerebral
Há um período inicial que vai até os 2-3 anos de idade, considerado o momento ótimo para a aquisição da linguagem.	O período que vai do nascimento até os 2-3 anos de idade é caracterizado pelo crescimento acelerado do cérebro, que passa de cerca de 30% de seus valores adultos para 70-80% desses valores.
Por volta dos 4 anos, o sistema lingüístico adquirido atinge um grau de organização mais ou menos estável, que pode ainda ser elaborado (e restaurado) até o início da puberdade.	A partir dos 3-4 anos, a taxa de crescimento desacelera gradualmente, e o cérebro progride contínua, mas lentamente até o começo da puberdade, em torno dos 12-13 anos.
A partir da puberdade, a aquisição primária da linguagem parece impossível, e a restauração do sistema adquirido é bastante restrita.	Nesta idade o cérebro chega, praticamente, à maturidade (ultrapassa 95% dos valores adultos): seu crescimento pouco progride até o fim da puberdade, quando atinge seu estado físico final.

Quadro 2. Correlações entre etapas do desenvolvimento lingüístico e do desenvolvimento cerebral. Adaptado de Lenneberg, 1967.

Menuzzi argumenta que a demonstração de uma interação sistemática entre o processo de maturação cerebral no homem e a manifestação de aspectos específicos da GU ofereceria um argumento decisivo tanto para a existência desses universais em um estágio inicial (S_0) quanto para a própria GU e a hipótese inatista.

Eis a relevância, portanto, da hipótese da maturação: se confirmada empiricamente, ela não apenas pode fornecer uma explicação para certos estágios de desenvolvimento da linguagem na criança, mas também pode sustentar uma das hipóteses mais controvertidas da lingüística e da psicologia cognitiva contemporâneas – a hipótese do inatismo, i. e., de que o homem é dotado biologicamente de capacidades mentais específicas para o desenvolvimento da linguagem. A questão fundamental, então, é: existe evidência satisfatória para a hipótese da maturação? (Menuzzi, 2001)

A busca por evidências satisfatórias para uma hipótese como esta – que implica investigar a relação entre a GU e as estruturas físicas do cérebro – escapa, no entanto, aos objetivos do programa gerativista (Raposo, 1992). As hipóteses da aprendizagem, da maturação cognitiva e da maturação lingüística continuam em franco debate, sem que haja um consenso quanto à questão de como a criança supera os estágios iniciais e reformula sua gramática provisória. Mais do que nunca, parece hoje inevitável pensar que parte dessa resposta esteja em processos biológicos característicos do desenvolvimento infantil. Dado o crescente acúmulo de conhecimento nessa área, parece ainda mais inevitável olhar para essas evidências em busca de restrições biológicas empíricas que auxiliem na delimitação de modelos teóricos lingüísticos.

O órgão da linguagem

A idéia de que o desenvolvimento de um “órgão mental” da linguagem seria responsável, ao menos em parte, pelo processo de aquisição foi originalmente sugerida por Lenneberg e, posteriormente, retomada por Chomsky. A proposição de um órgão para a

linguagem fundamenta-se em quatro características fundamentais que esta compartilha com outras habilidades (como caminhar, por exemplo) cuja emergência e o desenvolvimento são regulados pela maturação biológica do organismo, a saber:

- a. regularidade na seqüência de surgimento de determinados elementos, todos correlacionados com a idade e com outros marcos concomitantes do desenvolvimento;
- b. evidência de que as oportunidades de estimulação ambiental relevantes permanecem relativamente constantes ao longo do desenvolvimento (não podendo, portanto, explicar por si só os eventos), mas que a criança faz uso diferente de tais oportunidades à medida que cresce;
- c. observação de que as primeiras manifestações do comportamento – em parte ou completamente – emergem antes de que ele seja de utilidade imediata para o indivíduo;
- d. evidência de que as primeiras manifestações do comportamento não sinalizam uma prática dirigida a um objetivo.

A essas quatro características (a–d) comuns a formas de comportamento cuja emergência é maturacionalmente controlada, Lenneberg acrescentou um amplo leque de evidências anatômicas, morfológicas, fisiológicas, patológicas e evolutivas para sustentar sua tese de que a aquisição da linguagem é um processo dependente do desenvolvimento cerebral, com a conseqüente maturação do “órgão da linguagem” – e não o resultado de mudanças provocadas pelo ambiente. Como aponta Menuzzi (2001), seu objetivo, em última análise, era encontrar um argumento positivo em favor da hipótese da maturação; um argumento que não apenas mostrasse ser essa a hipótese mais plausível para explicar a seqüência de estágios lingüísticos observados no curso da aquisição, mas que também evidenciasse que essa seqüência corresponde, de modo mais ou menos preciso, a eventos específicos no

desenvolvimento anatomofisiológico do cérebro, os quais estariam regulando o surgimento de certas habilidades lingüísticas.

Apesar dos significativos avanços técnico-científicos observados neste início de século, o sonho de Lenneberg ainda parece algo distante. A observação simultânea de uma associação específica e direta entre o desenvolvimento de certas estruturas cerebrais – constituintes do assim chamado “órgão da linguagem” – e o aparecimento de construções particulares na fala infantil, embora teoricamente factível, ainda esbarra em limites intransponíveis. Apesar de ter avançado muito nos últimos anos, impulsionado principalmente pelas modernas técnicas de neuroimagem hoje disponíveis, o conhecimento acumulado sobre a organogênese cerebral ainda é bastante incompleto, em parte porque inevitavelmente questões éticas interpõem-se à realização de muitos estudos sobre a morfologia e a fisiologia do desenvolvimento normal do cérebro infantil.

Por outro lado, talvez nem o próprio Lenneberg houvesse sonhado com o exponencial avanço alcançado pela genética nas últimas décadas. Ao sustentar há meio século que a linguagem deveria estar, ao menos em parte, codificada em nossos genes, Lenneberg chamava a atenção para o fato de que nenhuma outra espécie – independentemente do nível de organização de sua estrutura social e cultural – compartilha com os seres humanos a habilidade lingüística em toda a sua complexidade. O próprio Lenneberg observou ainda que estudos com pares de gêmeos e genealogias familiares evidenciavam uma transmissão hereditária de distúrbios de linguagem – mas sabia que falar sobre “genes da linguagem” não era mais do que uma especulação, capaz de levantar “suposições duvidosas”, embora claramente entreviesse aí uma possibilidade de diálogo teórico entre dois campos tão distintos (1967:277).

A idéia de que exista uma Gramática Universal inata, geneticamente especificada, dando suporte ao processo de aquisição da linguagem por parte da criança, foi desenvolvida

muito antes da era da genômica e, talvez por isso, como comenta Jackendoff (2002), precise de uma nova roupagem e de alguns ajustes para se adequar aos tempos atuais. Rejeitada e duramente combatida por algumas correntes da lingüística, da psicologia e das neurociências, a hipótese da GU permanece digna da atenção que recebe há décadas e segue como ponto de referência para qualquer discussão sobre as relações entre genes e linguagem.

Um dos objetivos centrais da pesquisa em aquisição da linguagem ao longo dos anos tem sido justamente determinar os conhecimentos iniciais da criança no momento em que começa o processo de aquisição². A premissa básica é a de que deve haver alguma forma de conhecimento inato para justificar a rapidez e a relativa facilidade com que ocorre o desenvolvimento gramatical, especialmente quando comparado a outras habilidades cognitivas, que demandam instrução explícita (como a matemática, p.ex.).

Dentro do paradigma gerativista de investigação, padrões de linguagem observados consistentemente na fala de diversas crianças, em diferentes línguas, tornam-se potenciais candidatos a pertencer à GU, sendo interpretados, de modo geral, como indícios de algum tipo de conhecimento lingüístico inato. Alternativamente, esses padrões podem ser pensados como resultantes de uma estratégia geral de aprendizagem, dentro de uma abordagem própria da psicologia cognitivista. Embora essas duas leituras do processo diverjam em suas perspectivas teóricas, elas não são radicalmente inconciliáveis, já que é possível que ambas estejam parcialmente corretas.

O argumento de Chomsky em torno da GU (1965, 1981) não afirma que a globalidade do conhecimento lingüístico é inata, que a gramática de uma língua em particular é inata ou mesmo que o arcabouço estrutural das línguas é inato. O que é inato, para Chomsky, é uma base para o desenvolvimento da linguagem, um esquema inicial geral que gradualmente se

² O ponto de início do processo de aquisição é tradicionalmente entendido com sendo ao nascimento. Contudo, a possibilidade de que o desenvolvimento de habilidades lingüísticas tenha início ainda no ambiente intra-uterino vem sendo apontada por alguns pesquisadores, face à observação de diferentes capacidades discriminatórias acústicas. Ver os trabalhos do grupo de Jaques Mehler, incluindo Peña et al., 2003.

específica, à medida que a criança constrói a gramática da sua língua. O argumento, aqui, é contrário à *tabula rasa* – e não ao papel do ambiente ou à importância dos dados do *input* no processo de aquisição. Em uma passagem de 1965, Chomsky comenta a impossibilidade de caracterizar a composição desse esquema inicial:

What are the initial assumptions concerning the nature of language that the child brings to language learning, and how detailed and specific is the innate schema (the general definition of “grammar”) that gradually becomes more explicit and differentiated as the child learns the language? For the present, we cannot come at all close to making a hypothesis about innate schemata that is rich, detailed, and specific enough to account for the fact of language acquisition (Chomsky, 1965:27-8).

A composição específica da GU – as pré-especificações ou ferramentas das quais a criança se valeria para conseguir, em cerca de cinco anos, dominar a gramática da comunidade em que vive – permanece desconhecida e alvo de discussões. Em boa medida isso se justifica não apenas pela dificuldade em descobrir quais são esses elementos, ao mesmo tempo básicos e imprescindíveis para o desenvolvimento da linguagem, mas também porque a realidade desses elementos precisa ser admissível do ponto de vista biológico. É sobre o conteúdo da GU que recaem a maior parte das críticas sobre essa hipótese, especialmente porque os atributos investigados pelos lingüistas como possivelmente integrantes dessa gramática inicial são, em geral, específicos à linguagem e não mostram relação com outros domínios cognitivos conhecidos ou estudados.

De fato, como critica Tomasello (1995), as estruturas propostas por muitos lingüistas como pertencentes à GU são absolutamente desconhecidas e estranhas a qualquer não-lingüista – princípio de projeção, princípio da categoria vazia, princípio da subjacência. É indiscutível que tais princípios, como formulados pelos lingüistas, pressupõem a existência de funções altamente específicas e, nesse sentido, com base nos atuais conhecimentos sobre biologia molecular, é pouco provável que possam ser especificamente determinados por

genes. Contudo, o que parece escapar a essas críticas é a idéia de que esses princípios abstratos, específicos para a linguagem, possam ser, na verdade, desdobramentos do funcionamento de estruturas neurais especializadas, estas sim, geneticamente determinadas.

Em 1965, quando a pesquisa na área de aquisição ainda dava seus primeiros passos, Chomsky falava de universais formais e universais substantivos como sendo as duas categorias que compõem o repertório básico da GU. Os universais formais seriam basicamente um conjunto de regras utilizadas para combinar unidades básicas em estruturas complexas, incluindo também as regras restritivas da língua (*language constraints*). Os universais substantivos, por sua vez, compreenderiam elementos-chave para a construção da língua, como a capacidade de discriminar traços fonológicos, de perceber a segmentação das unidades de fala e de dominar a árvore sintática.

O conhecimento sintático é, indubitavelmente, o pomo da discórdia no que tange às pré-especificações da GU. Ao contrário dos demais elementos previstos por Chomsky (1965), que parecem poder depender de propriedades mais gerais, como propriedades combinatórias e acústico-discriminatórias, o conhecimento de regras restritivas e o domínio da estrutura sintática são domínios exclusivamente lingüísticos e altamente especializados. Considere um exemplo, apresentado por Caplan & Chomsky (1980). O que se deve notar aqui é que o pronome (*he* ou *him*) pode ser interpretado como fazendo referência ao sujeito *John* da oração principal em 1 e 4, mas não em 2 e 3.

1. *John believes he is incompetent.*
2. *John believes him to be incompetent.*
3. *John wants him to win.*
4. *John wants Bill to see him.*

Como ninguém é treinado, i.e., recebe instrução explícita para aprender essa regra, fica a questão de como ela é adquirida. A resposta, segundo a hipótese da GU, é a de que existe

uma faculdade da mente dotada de princípios específicos para lidar com a língua (neste caso em particular, com as propriedades de co-referência de diferentes tipos de nomes e pronomes). Princípios abstratos específicos para a linguagem guiarão o desdobramento do saber gramatical nos seres humanos, de modo que a gramática final do adulto contenha as especificações necessárias para compreender, distinguir e manejar as sutilezas das línguas humanas.

Se há mesmo um conjunto de pré-especificações dos quais a criança se vale para dar início à aquisição de sua língua – um conjunto de conhecimentos que é adquirido e não aprendido – é preciso explicar como a criança o adquiriu. A única explicação parece ser através do desenvolvimento de uma arquitetura cerebral específica, determinada a partir de uma combinação de herança genética e processos biológicos resultantes da expressão desses genes, expressão esta, por sua vez, guiada por uma combinação particular de fatores internos e externos (ambientais).

Apesar dos significativos avanços na compreensão do funcionamento do cérebro humano nas últimas décadas, ainda não foi possível elucidar exatamente quais são os mecanismos genéticos e neurais que subjazem a um desenvolvimento cerebral normal, nem mesmo discriminar precisamente quais as etapas e estruturas desse processo que são modificadas por distúrbios do desenvolvimento, nos quais se inclui a maioria dos distúrbios que afetam a linguagem. A descoberta de genes, como o *FOXP2*, oferece, contudo, uma via alternativa de investigação, permitindo explorar, a partir de uma perspectiva molecular, vias neurais relevantes ao desenvolvimento da fala e da linguagem.

Embora este trabalho esteja focado nas implicações da descoberta do *FOXP2*, não é apenas a descoberta de genes específicos que pode trazer contribuições valiosas para a discussão das relações entre cérebro e linguagem. Síndromes genéticas nas quais os indivíduos afetados apresentam alterações linguísticas específicas, com assimetria entre os

déficits observados, são também fontes valiosas de evidência sobre a organização e o desenvolvimento da linguagem e sua relação com os demais sistemas cognitivos, assim como são os casos de afasia, dislexia e autismo, que há mais tempo vêm atraindo a atenção dos pesquisadores.

A integração de dados moleculares com dados obtidos através de outras abordagens – como estudos de neuroimagem, testes lingüísticos e testes neuropsicológicos – pode fornecer novas evidências sobre estruturas cerebrais envolvidas no funcionamento da linguagem. Cada uma dessas evidências deve ser recebida com atenção e entusiasmo, sem perder de vista, contudo, que, no estado atual do conhecimento, cada uma dessas descobertas é tão somente uma pequena peça em um imenso quebra-cabeça.

CAPÍTULO II

DOS GENES À LINGUAGEM

Com quantos genes se faz um ser humano?

A primeira década do século XXI foi batizada como a do “despertar da genética cognitiva” (Pinker, 2001). O seqüenciamento do genoma humano e o avanço nas tecnologias de análise genética levaram os pesquisadores a investir na possibilidade de, finalmente, trazer à luz a base daquilo que nos faz verdadeiramente humanos – nossa cognição. A pergunta, então, era a seguinte: por tudo aquilo que se sabe, é possível que a função de genes específicos esteja diretamente ligada a funções cognitivas como pensamento, memória, cálculo e linguagem?

A resposta para essa pergunta passa essencialmente por um conhecimento preliminar acerca do papel desempenhado pelos genes no desenvolvimento humano. Durante muito tempo se acreditou na idéia de que a complexidade de um organismo seria diretamente proporcional à sua quantidade de material genético, i.e., ao número de genes de cada espécie. Com base naquilo que se sabia sobre o genoma de outras espécies, estimava-se que o genoma humano tivesse cerca de 100 mil genes. Hoje, no entanto, sabemos que ele é composto por aproximadamente 30.000 genes – um número muito similar ao de camundongos, cachorros, macacos ou qualquer outro mamífero (e inferior ao de muitas plantas). Não está, portanto, no número de genes a maior ou menor complexidade de cada uma das espécies.

Nossos genes também não são, em si, muito distintos dos genes de outras espécies. Com os chimpanzés, por exemplo, compartilhamos aproximadamente 95% de nosso DNA (Britten, 2002), fazendo com que a identidade genética entre homem e chimpanzé seja muito

maior do que aquela entre chimpanzés e gorilas. Embora evolutivamente os chimpanzés constituam a espécie mais próxima ao homem – apresentando uma rica organização social e sendo capazes de reconhecer a própria imagem no espelho, utilizar ferramentas e aprender alguns sinais de comunicação – são obviamente muitas as diferenças entre as duas espécies. Como, então, a comparação entre o genoma de homens e chimpanzés revela uma diferença tão pequena?

Parte dessa resposta pode ser pensada a partir de uma analogia literária: a frase que abre o *Ano da Morte de Ricardo Reis*, de José Saramago, diz “Aqui o mar acaba e a terra principia”, invertendo o verso camoniano de *Os Lusíadas* “Donde a terra se acaba e o mar começa”. Embora sejam duas obras completamente distintas, é certo que Saramago e Camões utilizam ao longo de suas narrativas centenas de palavras em comum. Existem palavras que Saramago usa, mas não Camões, como *hotel*, *trem* e *telefone*; e palavras que Camões usa, mas não Saramago, como *superbíssimo*, *hórridos* e *denodadas*. No entanto, essas serão poucas comparadas com a quantidade de palavras em comum. Logo, o fundamento da diferença entre essas obras não está no uso de um conjunto diferente de palavras, mas no mesmo conjunto de palavras usadas em padrão e forma totalmente distintos. Do mesmo modo, a diferença entre o homem e o chimpanzé não está no uso de genes diferentes, mas nos mesmos genes usados de forma diferente (analogia adaptada de Ridley, 2003).

Como nas demais espécies, os 30.000 genes humanos estão distribuídos nos cromossomos, que por sua vez estão contidos no núcleo de cada célula do organismo (Fig. 2). Na espécie humana, são 46 cromossomos, organizados aos pares, presentes em cada um dos trilhões de células do corpo: 23 são herdados da mãe (cromossomos maternos) e 23, do pai (cromossomos paternos). A exceção fica por conta das células germinativas – óvulos e espermatozoides – que contêm um único conjunto de 23 cromossomos, que irão se fundir no momento da fecundação, restabelecendo o número inicial.

Dentre os 46 cromossomos, 44 são chamados cromossomos autossômicos, ou autossomos, constituindo pares semelhantes entre si em forma e conteúdo – daí a expressão 22 pares de “cromossomos homólogos”. O último par, por sua vez, corresponde aos dois cromossomos sexuais, X e Y, que apresentam morfologia e material genético diferente entre si e determinam o sexo do indivíduo: enquanto as mulheres apresentam dois cromossomos X (44 autossomos + XX), os homens têm um cromossomo X e outro Y (44 autossomos + XY).

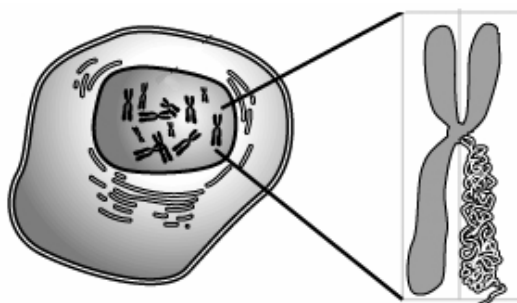


Figura 2. No núcleo de cada célula somática do organismo humano, encontram-se 46 pares de cromossomos (apenas alguns cromossomos estão ilustrados na figura).

Desde a década de 1950, já se conhecem técnicas simples e eficazes para a observação dos cromossomos humanos a partir de células extraídas do sangue periférico. Os 46 cromossomos, nos quais estão contidos os genes, podem ser visualizados ao microscópio óptico durante um estágio específico da divisão celular, a metáfase, quando se encontram mais fortemente condensados. O número de cromossomos varia de acordo com a espécie, e o conjunto cromossômico típico de cada espécie é chamado cariótipo. A Fig. 3 representa o cariótipo humano normal de um indivíduo do sexo masculino (44 + XY).

O fato de que o conjunto de genes humanos encontra-se distribuído ao longo dos 23 cromossomos implica que, na ausência de mutações ou outros fenômenos complexos, cada indivíduo terá, via de regra, duas cópias do mesmo gene, uma localizada no cromossomo materno e outra na região homóloga do cromossomo paterno (diz-se, assim, que cada gene possui um *locus* situado em regiões homólogas de ambos os cromossomos). Esta é uma característica importante na determinação da manifestação dos distúrbios genéticos, uma vez

que alguns serão observados quando apenas um dos genes estiver mutado (distúrbios dominantes), enquanto outros se manifestarão somente quando ambas as cópias, paterna e materna, estiverem afetadas (distúrbios recessivos).

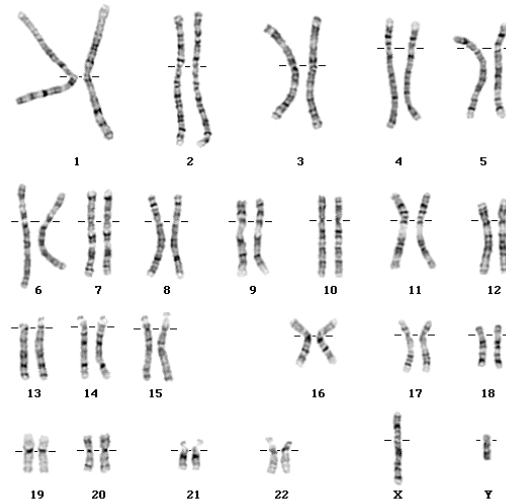


Figura 3. Cariótipo humano, mostrando 22 pares de cromossomos autossômicos e um par de cromossomos sexuais. A presença dos cromossomos X e Y no cariótipo indica sexo masculino.

Todavia, é importante ressaltar desde aqui que a manifestação de doenças a partir de mutações em um único gene é um evento relativamente raro. Sabe-se hoje que muitas das doenças de maior prevalência na população – como câncer, doenças coronarianas, diabetes, hipertensão, depressão e asma – têm sua etiologia, em alguma medida, relacionada a um componente genético, sendo resultantes da ação conjunta de diversos genes em interação com determinados estímulos ambientais. Nesses casos, cada um dos genes envolvidos na etiologia tem um efeito pequeno sobre a manifestação clínica da doença. De modo diverso, um número relativamente pequeno de síndromes é resultado de mutações em um único gene: observam-se, então, fenômenos monogênicos – como fibrose cística, distrofia muscular de Duchenne, doença de Huntington e fenilcetonúria, por exemplo –, que têm como característica o fato de que o gene mutado tem um grande efeito, acarretando sozinho conseqüências desastrosas para a saúde do indivíduo.

Ainda sobre a organização do material genético, de modo geral, pode-se afirmar que cromossomos maiores (como os cromossomos 1, 2 e 3) apresentam um número maior de genes, enquanto há uma quantidade menor de genes em cromossomos pequenos, como o 21, o 22 e o Y (ver Fig. 3). No entanto, o fato de essa correlação não ser uma regra traz à luz uma pergunta importante: se os cromossomos contêm o material genético na forma de genes e não há uma relação obrigatória entre o tamanho do cromossomo e o número de genes, o que mais há em um cromossomo?

Na verdade, os cromossomos representam um estado máximo de agregação, condensação ou enrolamento das moléculas de DNA (ácido desoxirribonucléico), a chamada “molécula da vida”, na qual a informação genética de um organismo é armazenada e transmitida para seus descendentes. As moléculas de DNA, por sua vez, são compostas por três estruturas: um fosfato, um açúcar e uma base nitrogenada. Em conjunto, esses três elementos formam um nucleotídeo. Como o fosfato e o açúcar (a desoxirribose) são invariáveis, são as diferentes bases nitrogenadas que determinam os quatro diferentes tipos de nucleotídeos existentes no DNA: adenina, timina, guanina e citosina (representados pelas letras A, T, G e C) (Fig. 4). Os nucleotídeos unem-se uns aos outros, formando uma cadeia ou uma “fita”, que é complementada por uma segunda fita de nucleotídeos, que obedece uma ordem rígida de pareamento entre as bases: A só pareia com T, C só com G (e vice-versa).

Ao todo, o genoma humano contém cerca de três bilhões de pares de nucleotídeos (IHGSC, 2001; Venter et al. 2001), distribuídos ao longo dos cromossomos. O DNA é, portanto, um código linear de três bilhões de letras seqüenciadas de modo específico. Se fosse possível esticar a fita de DNA, estima-se que ela teria 1,8 metros de comprimento (note, assim, que o grau de condensação da molécula de DNA é tamanho que o microscópico núcleo de uma célula consegue armazenar quase dois metros de pares de bases).

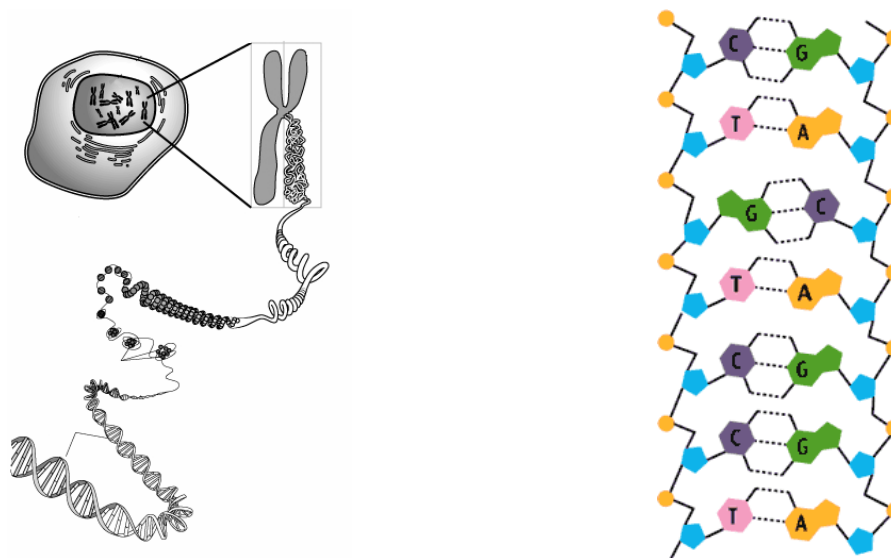


Figura 4. À esquerda: o cromossomo corresponde ao estado máximo de agregação da molécula de DNA. À direita: estrutura esquemática da molécula de DNA, organizada na forma de uma escada, com duas fitas interligadas, seguindo um pareamento específico. Os círculos nas laterais representam os fosfatos; os pentágonos representam o açúcar; e as quatro bases nitrogenadas mostram o padrão fixo de pareamento A-T, C-G. A descoberta da estrutura do DNA é atribuída a Watson e Crick (1953).

Os quatro pares de bases que formam o DNA (A, T, C e G) são comuns às diferentes espécies; assim, há uma relativa universalidade do código genético. Além disso, um dos mais importantes achados do Projeto Genoma Humano revelou a natureza da diversidade existente entre diferentes indivíduos de uma mesma espécie. Cada pessoa que já existiu no planeta – com exceção de gêmeos idênticos – possui um genoma único. Ainda assim, num aparente paradoxo, quaisquer dois genomas humanos são ~99,9% idênticos. Isso porque esse 0,1% de variação ainda deixa espaço para milhões de diferenças entre os 3,2 bilhões de pares de bases de nucleotídeos que compõem o genoma. De modo importante, o Projeto Genoma revelou que o nível de identidade entre dois genomas é aproximadamente o mesmo independentemente da origem étnica dos indivíduos, ou seja, dois indivíduos brancos europeus comparados entre si mostram o mesmo grau de diferença genética que o resultante da comparação entre um europeu e um asiático, minando, assim, com o conceito de raça (Ojopi, 2004).

Cada gene seqüenciado corresponde a um segmento específico da fita de DNA; cada gene é, portanto, uma longa seqüência única de ordenamento das quatro bases químicas. Existem seqüências de bases que formam genes, enquanto outras, a imensa maioria, não. O

que é peculiar a cada gene é que a sua seqüência é funcional: ela contém as informações básicas para a formação de alguma proteína necessária ao adequado funcionamento do organismo.

Dentre as surpresas ocasionadas pelo anúncio do seqüenciamento quase completo do genoma humano, em 2001, uma foi particularmente relevante. A porção claramente funcional de DNA (i.e., aquela que dá origem a genes) é muito pequena, menos de 3% do genoma. Ou seja, dos três bilhões de pares de bases, apenas uma diminuta fração faz parte, de fato, de algum dos 30.000 genes. A maior parte das bases tem função desconhecida e não participa da composição de genes, compondo o chamado DNA intergênico.

O fato de ser tão pequena a parcela de DNA que constitui genes lançou luz sobre a necessidade de investigar o papel do material presente no entorno dos genes, o qual foi inicialmente considerado como “DNA lixo” – seqüências que se acumularam por milhões de anos sem utilidade conhecida. Atualmente diversas funções para esse DNA não-codificador (em torno de 97% do genoma humano) vêm sendo investigadas, especialmente no sentido de que esse material contenha seqüências regulatórias capazes de controlar o nível de expressão dos diferentes genes, entre outras funções.

Em síntese, um gene é uma seqüência específica de DNA (ou seja, uma seqüência específica de nucleotídeos) que tem a propriedade de conter a informação necessária para codificar uma proteína. Assim, considerando um cromossomo qualquer, o cromossomo 7, por exemplo, onde se situa o gene *FOXP2*, pode-se imaginar que, em uma determinada porção desse cromossomo, estará um gene *X*. Esse gene *X* é, na verdade, uma seqüência de DNA composta por um ordenamento específico de centenas ou milhares de nucleotídeos em cadeia (para se ter uma idéia, o menor gene humano tem 252 pares de base; o maior tem mais de um milhão de repetições A, T, C, G). O que faz dessa seqüência um gene, além de sua organização estrutural específica, é sua capacidade única de dar origem à proteína hipotética

(pX). Ao lado dessa seqüência pode haver centenas pares de bases que não codificam proteínas, até que mais uma centena de nucleotídeos à frente apareça um novo gene Z, que forma a proteína pZ.

Uma analogia bastante comum equipara as seqüências de bases do DNA a uma receita de bolo para a formação de uma proteína. Note que o mais importante é o bolo, não a receita, mas há uma relação de interdependência entre ambos, no sentido de que o DNA é o livro que contém a informação para dar origem a um produto, no caso, uma determinada proteína.

If molecular biologists had to designate one category of macromolecules as being essential for life, it would be proteins and their multiple functions, not DNA and genes. Genes are important only because they contain enough information to enable the synthesis of these proteins at an appropriate time and place (Morange, 2001:2).

Outro ponto importante a destacar é que a molécula de DNA não é capaz de deixar o núcleo da célula e, por isso, a informação armazenada no DNA precisa ser lida por uma molécula intermediária, neste caso, o RNA (ácido ribonucléico). Diz-se, assim, que o DNA é capaz de ser transcrito em RNA, esta sim uma molécula capaz de deixar o núcleo da célula e migrar em direção ao citoplasma para, então, dar origem a uma proteína. Nessa última etapa do processo, o RNA, originado a partir de uma seqüência específica de DNA, é traduzido em uma proteína. Este é o chamado dogma central da biologia molecular: DNA → RNA → proteína.

Além das seqüências de bases que se situam entre dois genes, o chamado material intergênico, mesmo no interior de um gene, que é uma seqüência tipicamente codificadora de proteína, há pares de base que não farão parte do produto protéico formado. No processo de transcrição do DNA para o RNA, eles serão removidos, através de um processo chamado *splicing*, e não irão participar da formação final da proteína (Fig. 5). Esses segmentos não-

codificadores são chamados íntrons (seqüências intercalares); já as partes do gene que, de fato, entram na formação final da proteína correspondem aos éxons (seqüências codificantes).

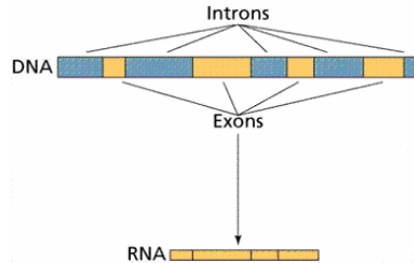


Figura 5. Estrutura representativa de um gene humano, composto por seqüências de íntrons e éxons; no processo de *splicing*, os íntrons são removidos e apenas as bases que compõem os éxons entram na formação da proteína, a partir do RNA.

DNA e RNA são compostos essencialmente pelos mesmos elementos estruturais, os nucleotídeos (com a exceção de que no RNA não há a base timina e, em seu lugar, aparece uma uracila, modificando o pareamento para A-U; C-G). Já os componentes estruturais das proteínas são os aminoácidos; cada grupo de três bases seqüenciais do RNA está relacionado a um determinado aminoácido (Fig. 6). Existem apenas 21 aminoácidos, que, combinados nas mais diversas seqüências, dão origem a milhões de proteínas. Assim, uma seqüência de nucleotídeos do DNA é transcrita em uma seqüência complementar de nucleotídeos do RNA, que é “lida”, sempre de três em três bases, determinando a seqüência de aminoácidos que formará a proteína.

		SEGUNDA LETRA				
		U	C	A	G	
PRIMEIRA LETRA	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu	Ser	Stop	(Stop) Trp	A
		Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U	
	Leu	Pro	His	Arg	C	
	Leu	Pro	Gin	Arg	A	
	Leu	Pro	Gin	Arg	G	
A	Ile (Met)	Thr	Asn	Ser	U	
	Ile	Thr	Asn	Ser	C	
	(Ile) Met	Thr	Lys	(Arg) Stop	A	
	Ile	Thr	Lys	(Arg) Stop	G	
G	Val	Ala	Asp	Gly	U	
	Val	Ala	Asp	Gly	C	
	Val	Ala	Glu	Gly	A	
	Val	Ala	Glu	Gly	G	

Figura 6. Organização do código genético. Cada três pares de base no RNA (letras) correspondem a um dos 21 aminoácidos conhecidos (note que diferentes trinças podem formar um mesmo aminoácido: p. ex., CGU e CGC são duas das seqüências que podem gerar Arg – arginina). É a união de vários aminoácidos que dá origem às proteínas.

Cabe ressaltar, contudo, que ao contrário do que se acreditava por muito tempo, um gene não necessariamente dá origem a uma única proteína – o que derruba a crença na díade “um gene – uma proteína”. Mecanismos ainda pouco conhecidos permitem que uma mesma seqüência de nucleotídeos no DNA dê origem a diferentes seqüências de aminoácidos, originando, portanto, diferentes proteínas. Isso ocorre principalmente através de um fenômeno conhecido como *splicing* alternativo, no qual variações na clivagem da molécula de RNA permitem arranjos combinatórios de diferentes éxons (Fig. 7). Isso reforça a noção de que a complexidade de uma espécie não pode ser antecipada pelo número total de genes contidos em seu genoma; no caso da espécie humana, estima-se que mais de 100 mil proteínas possam ser derivadas do conjunto de pouco mais de 30.000 genes (IHGSC, 2001).

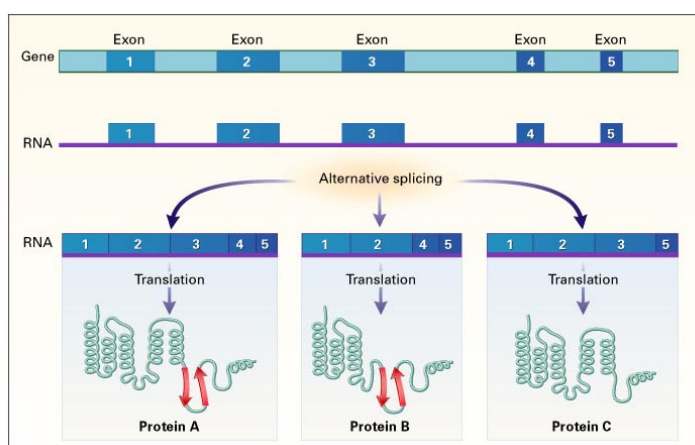


Figura 7. Mecanismo de *splicing* alternativo, através do qual uma única molécula de RNA apresenta diversas possibilidades de processamento (combinação de éxons), ampliando o número de proteínas diferentes sintetizadas a partir de uma mesma seqüência original de DNA (extraído de Gutmacher & Collins, 2002).

Além da função básica do DNA de armazenar a informação necessária à promoção da síntese de proteínas, de modo bastante importante, cabe também ao DNA um papel evolutivo. Ao garantir o armazenamento adequado da informação genética, essas moléculas permitem o repasse desse material através das gerações, dando à prole a oportunidade de produzir as mesmas proteínas que permitiram aos seus progenitores adaptar-se com sucesso ao meio em que vivem.

Genes e proteínas orquestrando o desenvolvimento

O DNA e, conseqüentemente, o conjunto de genes de cada célula somática é, essencialmente, o mesmo. O DNA de uma célula do fígado, um hepatócito, não difere do DNA de uma célula nervosa, um neurônio, que por sua vez não difere do de uma célula cardíaca. Todas tiveram origem a partir de uma única célula, o zigoto, resultante da fusão original do espermatozóide com o óvulo, unindo o material genético contido em 23 cromossomos paternos e 23 maternos. O que permite a essas células tornarem-se tão diferentes entre si – morfologicamente e, por conseqüência, funcionalmente – são os diferentes genes expressos por cada uma delas, em diferentes momentos ao longo do seu ciclo de vida (Fig. 8). O fato de que cada tipo celular é capaz de produzir um conjunto específico de proteínas estabelece uma das principais distinções entre organismos simples e complexos.

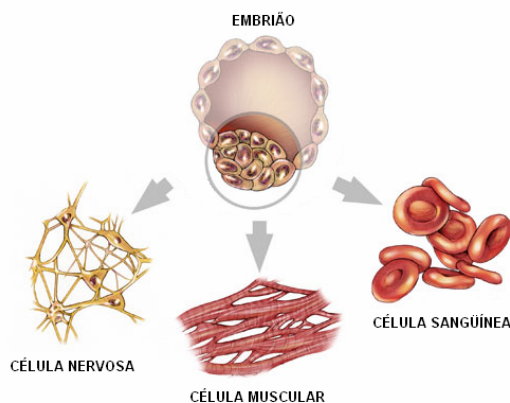


Figura 8. Após a formação do zigoto, os processos iniciais de divisão celular dão origem a um pequeno conjunto de células embrionárias, não-especializadas e multipotentes, i.e., capazes de dar origem a diversos tipos celulares, do qual se originam as demais células do organismo, com formas e funções distintas.

A cada subdivisão da célula original (zigoto) observa-se um nível crescente de especialização, com as células recém-formadas passando de um estado totalmente indiferenciado, no qual podem vir a integrar qualquer tecido, para um estado diferenciado, quando assumem a forma e a função de um tipo celular específico.

Ainda não se compreende inteiramente o processo de diferenciação celular. Sabe-se, contudo, que ele depende de sinais provenientes de hormônios, da matriz extracelular que se

forma no entorno das células, do próprio contato direto entre células e de fatores de diferenciação, conhecidos como citocinas. Já no início do desenvolvimento do organismo, nos estágios iniciais da embriogênese, observa-se um certo grau de compartimentalização celular, quando grupos de células começam a se posicionar de acordo com a função que irão desempenhar mais tarde. A orquestração dos sinais recebidos pela célula resulta, então, na repressão de certos genes e na ativação de outros, alterando a forma da célula, seus produtos, sua própria estrutura e as moléculas de sua superfície, essenciais para a comunicação intercelular. Assim, seqüências de modificações bioquímicas, morfológicas e funcionais transformam uma célula primitiva indiferenciada, que executa diversas funções com baixa eficiência, em uma célula especializada, capaz de executar uma função determinada com alta eficiência (Junqueira & Carneiro, 1999).

Foi a partir de estudos com a mosca-das-frutas (*Drosophila melanogaster*) – um inseto muito freqüentemente empregado em pesquisas de biologia molecular – que um grupo de geneticistas descobriu, ainda na década de 1980, um importante conjunto de genes diretamente implicados na coordenação do desenvolvimento embrionário. Christiane Nüsslein-Volhard e Eric Wieschaus identificaram 15 genes-chave na determinação do padrão de distribuição dos segmentos corporais da mosca-das-frutas, que é composta por cabeça, abdome, tórax e cauda. Um dos genes, por exemplo, é mais atuante no início da embriogênese e determina a polaridade do animal (o eixo cabeça-cauda), ou seja, onde deve ser formada a cabeça e onde deve surgir a cauda. Paralelamente, Edward Lewis descreveu os chamados genes homeóticos, que atuam mais tarde, regulando o desenvolvimento de um segmento larval em um segmento corporal específico. Os genes homeóticos mais estudados receberam os sugestivos nomes de complexo *Bitórax* e complexo *Antennapedia*, por terem sido estudados em animais mutantes que desenvolveram um par extra de asas (onde normalmente deveriam estar os balancins da mosca) e patas na cabeça, no lugar de antenas, respectivamente

(Alberts, 2002). Nüsslein-Volhard, Wieschaus e Lewis receberam, em conjunto, o prêmio Nobel de Medicina de 1995 por tais descobertas.

Anteriormente a essas descobertas, havia um certo ceticismo entre os embriologistas quanto à possibilidade de que os genes regessem o desenvolvimento inicial do organismo. A descoberta dos genes homeóticos alterou esse panorama, mostrando que alguns genes, cujos produtos protéicos são expressos num estágio bastante precoce do desenvolvimento, são capazes de determinar padrões básicos do crescimento embrionário, agindo sobre uma região definida de segmento corporal (Morange, 2001). Esses genes foram, posteriormente, agrupados numa mesma família, denominada *Hox*, e pesquisadores que investigavam a embriogênese de camundongos descobriram que genes *Hox* estavam presentes não apenas na mosca-das-frutas, mas também nesses animais. Assim, por exemplo, os genes *Hox* com ação no local correspondente àquele da formação de asas na mosca, no camundongo, informavam o local (não o modo) adequado para o desenvolvimento de costelas nos roedores.

A particularidade dos genes *Hox* diz respeito ao seu modo de ação. Os mais de 30 genes dessa família atualmente conhecidos têm em comum uma seqüência de bases de DNA característica, chamada *homeobox* (daí a designação *Hox*). A seqüência *homeobox* é reconhecida por proteínas – chamadas fatores de transcrição – que se ligam a essa região (conhecida como *DNA-binding domain* ou domínio de ligação ao DNA) e ativam a transcrição desses genes. Os genes *Hox*, por sua vez, também dão origem a proteínas do tipo fatores de transcrição, que irão atuar sobre outros genes, promovendo sua ativação ou repressão, em um grande efeito dominó (Fig. 9).

Para entender o papel dos fatores de transcrição é válido pensar nos genes como sendo constituídos por dois elementos-chave: uma região codificadora e uma região reguladora. A região codificadora de um gene é aquela que contém todas as informações necessárias para a síntese da proteína, contendo uma seqüência específica de aminoácidos. Como discutido

anteriormente, para a formação da proteína, a seqüência de DNA de cada gene é transcrita em cópias intermediárias – as moléculas de RNA. O papel da região reguladora é, entre outras coisas, determinar, a cada momento, quantas cópias desse RNA serão feitas. Fatores de transcrição são proteínas que interagem com essas regiões reguladoras dos genes, modulando o nível de expressão de outras proteínas e influenciando, em última análise, a abundância dessas proteínas nas células (Marcus & Fisher, 2003).

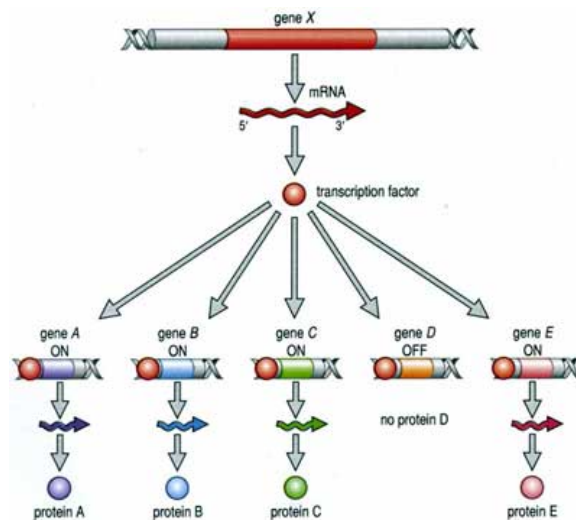


Figura 9. O gene *X* é transcrito em RNA e dá origem a uma proteína pertencente à classe dos fatores de transcrição. Essa proteína, por sua vez, irá se ligar a outros genes, ativando sua transcrição (genes A, B, C e E) ou bloqueando-a (gene D).

Os fatores de transcrição são, portanto, proteínas reguladoras que se ligam a uma região específica do DNA que, então, responde a essa ligação, transcrevendo ou não um conjunto particular de genes. Esse modo de ação faz com que esses genes estejam, por vias indiretas, relacionados a diferentes funções orgânicas, influenciando na regulação da síntese de proteínas que serão expressas em diferentes tecidos. Assim, quando um gene que é ele próprio o código para um fator de transcrição sofre alterações em sua expressão, a consequência é uma cascata de modificações na expressão de outros tantos genes (Ridley, 2003).

Não são apenas os genes *Hox*, mas diversos outros genes – incluindo os genes da família *FOX*, entre eles, o *FOXP2* – que agem controlando a expressão de outros genes. O termo expressão gênica refere-se ao processo pelo qual a informação codificada por um

determinado gene é decodificada em uma proteína. Os fatores de transcrição, ao agir como reguladores desse processo, podem levar a uma expressão gênica diferencial, permitindo que células que compartilham o mesmo material genético possam se diferenciar umas das outras pela seleção e controle das proteínas que serão produzidas. Além de fundamental para o adequado desenvolvimento embrionário, o controle da expressão gênica objetiva principalmente permitir que as células se ajustem às mudanças no ambiente, de forma a otimizar sua possibilidade de sobrevivência e desenvolvimento.

Nas mais diferentes espécies, a transcrição do DNA em RNA para, então, dar origem à proteína requer a participação de uma enzima, conhecida como RNA-polimerase. A RNA-polimerase bacteriana é capaz de iniciar a transcrição sem o auxílio de proteínas adicionais; contudo, as RNA-polimerases humanas requerem a colaboração de um grande conjunto de proteínas, os chamados fatores gerais de transcrição. Os fatores gerais de transcrição são proteínas essenciais para a transcrição de todos os genes; existem também os fatores específicos ou seletivos de transcrição, grupo no qual se inclui o *FOXP2*, que são necessários para a transcrição de apenas alguns genes. Ambos agem através da sua ligação a regiões específicas do DNA da célula, interferindo a partir daí em etapas essenciais do processo de transcrição (como, p.ex., a separação das fitas de DNA), de modo a reprimir ou ativar a transcrição de genes naquela célula.

Do ponto de vista evolutivo, acredita-se que boa parte das diferenças observadas entre as espécies origine-se de modificações relacionadas aos fatores de transcrição, muito embora a maioria desses genes encontre-se altamente conservada entre as espécies (Morange, 2001; Ridley, 2003). Mutações não apenas nos genes que codificam fatores de transcrição, mas principalmente nas regiões regulatórias às quais essas proteínas se ligam, podem ocasionar importantes variações na atividade gênica, por influenciar a expressão de um conjunto de genes.

Proteínas reguladoras, em geral, não agem independentemente. Existe um controle combinatorial sobre os genes humanos, no sentido de que eles são normalmente regulados por um conjunto de proteínas que trabalham juntas para determinar o nível de expressão de um único gene. A maioria das proteínas reguladoras de genes atua como parte de um “comitê” de proteínas, todas essenciais para a expressão de um determinado gene na célula correta, em resposta a uma dada condição, no tempo certo e no nível requerido.

Dessa forma, o destino de uma célula embrionária é determinado pela expressão gênica durante o período de desenvolvimento. A regulação desse mecanismo traz conseqüências não somente para a diferenciação celular (que tipo de célula será formada), mas também para a organização espacial do conjunto de células na formação dos tecidos (para onde as células devem migrar).

Após o seqüenciamento do genoma humano, os pesquisadores estão se valendo das informações já conhecidas sobre regiões reguladoras, genes e proteínas para agora se debruçar sobre a função específica de cada um desses genes no funcionamento do organismo. O atual conjunto de ferramentas disponíveis tornou possível investigar o conjunto de genes expressos em uma célula frente a uma condição fisiológica particular. Assim, é possível identificar os genes que são expressos ou reprimidos em determinadas situações biológicas, incluindo a possibilidade de identificar os aspectos temporais e espaciais da expressão de genes envolvidos no desenvolvimento.

No caso do sistema nervoso central, o processo de expressão gênica é fundamental para o estabelecimento de regiões especializadas em determinadas funções. Estudos recentes têm procurado estimar a contribuição genética para diferenças individuais na constituição de estruturas cerebrais. Thompson et al. (2001) demonstraram que fatores genéticos influenciam significativamente o volume neuronal de estruturas corticais, sendo a concordância entre gêmeos monozigóticos superior a 90% no que diz respeito a regiões do cérebro fortemente

ligadas ao desempenho lingüístico (áreas de Broca e Wernicke). O desafio que se coloca, então, é o de descobrir que genes são responsáveis por esse fenômeno e qual a sua participação no desenvolvimento de traços fenotípicos como a linguagem.

Em suma, genes – em especial os regulatórios – desempenham um papel essencial no desenvolvimento humano, fazendo com que o embrião sintetize produtos celulares específicos, em momentos determinados, permitindo, assim, a organização dos diferentes órgãos e sistemas. Todo esse processo é mediado não só pelas interações gene-gene, como no caso dos fatores de transcrição, mas também pelas interações genes-ambiente, nas quais estímulos externos ativam sinalizadores celulares que também contribuem para a regulação do processo de expressão gênica.

Moléculas da mente

A pergunta com a qual se inicia este capítulo pode agora ser retomada. Existem genes envolvidos no desenvolvimento de habilidades cognitivas humanas e, em particular, no desenvolvimento da linguagem? A existência de casos de “crianças selvagens”, como os famosos Genie e Victor – crianças isoladas de qualquer contato humano desde muito cedo na infância e que, ao serem descobertas, já em idade puberal, não conseguiram adquirir habilidades lingüísticas normais – evidencia o papel irrevogável da interação social no desenvolvimento da linguagem. Todavia, esses casos (e todo o corpo de evidências sobre o papel do *input* e da socialização) em nada contradizem a idéia de que genes sejam também essenciais ao processo de aquisição.

Apesar do grande desenvolvimento nas áreas de biologia molecular, neurociências e ciências cognitivas, ainda se sabe muito pouco sobre a natureza dos genes que nos fazem humanos. Contudo, as informações hoje disponíveis permitem discutir a plausibilidade de

certas proposições acerca das relações entre genes e linguagem. Um primeiro elemento a considerar nessa relação é que o caminho até a descoberta de um gene específico diretamente implicado em um dado fenótipo (i.e., uma característica observável, manifestada como resultado da interação entre a constituição genotípica do indivíduo e o ambiente em que vive) é longo e tortuoso – especialmente quando o fenótipo em questão é complexo, como são as habilidades cognitivas humanas.

A linguagem – assim como a memória, a inteligência e tantas outras funções cognitivas – pertence ao grupo das chamadas heranças complexas ou multifatoriais, isto é, características biológicas cuja expressão depende não de um único gene (como no caso das heranças simples ou monogênicas), mas de um conjunto de genes em associação a determinados estímulos ambientais. No caso da linguagem (e da ampla maioria dos fenótipos complexos) ainda não se conhecem quais são precisamente os genes que contribuem para sua manifestação, mas sabe-se que há mais de um e suspeita-se que sejam muitos.

Há apenas algumas décadas, o campo da psicologia estava dominado pela idéia de que o comportamento era decorrente da influência de fatores quase puramente ambientais; notadamente este fato se refletiu na lingüística e, em especial, na aquisição da linguagem, inicialmente com a abordagem behaviorista e, mais recentemente, com os modelos conexionistas de aprendizagem da língua (ver Elman, 1996). Hoje, no entanto, há um razoável consenso quanto à importante influência dos fatores genéticos sobre os mais diversos aspectos do comportamento humano e assim, naturalmente, sobre a etiologia de transtornos comportamentais, tais como a esquizofrenia, o alcoolismo, o autismo e o transtorno de déficit de atenção e hiperatividade (Sherman et al., 1997).

O que permanece em discussão hoje é qual a parcela dessa influência genética; uma resposta definitiva necessariamente deve passar pela descoberta de quais genes, de fato, contribuem para a emergência do comportamento investigado. Mais do que descobrir quais

são os genes, o grande desafio que se impõe hoje à genética cognitiva (e a descoberta do *FOXP2* irá ilustrar isso de maneira categórica) é desvendar o papel, a função desses genes no desenvolvimento e no funcionamento dos traços a eles relacionados.

O conhecimento acerca das bases genéticas de doenças complexas – como os transtornos de linguagem em geral – vem-se ampliando paulatinamente nos últimos anos, permitindo uma melhor compreensão sobre a etiologia de doenças neuropsiquiátricas e outras patologias de alta prevalência e impacto na sociedade, como diabetes, hipertensão e arteriosclerose. O seqüenciamento do genoma humano representou, sem dúvida, um passo essencial no entendimento da biologia humana, trazendo excelentes perspectivas de avanço para o campo de pesquisas biomédicas e incentivando a busca por novas abordagens para determinar os riscos associados a uma determinada doença ou o desenvolvimento de terapias individualizadas. Contudo, é importante notar que o seqüenciamento do genoma é apenas um começo, devendo tais informações genéticas serem cruzadas com dados clínicos, étnicos e ambientais para a busca de um melhor entendimento das bases biológicas dos fenótipos complexos (Ojopi et al., 2004).

Em 2001, o anúncio do seqüenciamento quase completo do genoma humano, finalizado em 2003, trouxe aos pesquisadores e à comunidade científica em geral a expectativa de ampliar ainda mais, e com maior rapidez, o entendimento das bases moleculares subjacentes a tais fenótipos. No entanto, o quadro que emergiu a partir de 2001 não confirmou essa expectativa, o que, em parte, ajuda a entender por que a descoberta do *FOXP2* foi saudada com grande entusiasmo. Embora amplo financiamento tenha sido colocado à disposição dos muitos centros de pesquisa genética voltados à investigação de doenças complexas em todo o mundo, poucos resultados conclusivos foram alcançados nesse campo. Nos Estados Unidos, mais de 1.300 projetos foram apoiados pelo *National Institutes of Health*, o poderoso órgão financiador de pesquisas em saúde do governo norte-americano,

para o estudo de mais de 170 genes supostamente envolvidos nas mais variadas heranças complexas. Os resultados foram, no mínimo, desalentadores: menos de 50 variações gênicas foram identificadas e provaram estar, de fato, associadas à manifestação das doenças pesquisadas. Mais do que isso, apenas 16% a 30% dos resultados desses estudos foram replicados consistentemente por trabalhos posteriores (Page et al., 2003).

Ressalte-se que esses dados pouco animadores dizem respeito a heranças complexas em geral, não estando restritos aos resultados obtidos no campo de pesquisa da genética do comportamento. Se assim fosse, presumivelmente, os resultados seriam ainda mais negativos, já que essa área inclui um fator complicador adicional que não pode ser desprezado: a caracterização do fenótipo. Como, metodologicamente, toda pesquisa requer, no mínimo, a identificação de sujeitos que componham o grupo dos “casos” é preciso delimitar com clareza as características que são comuns ao grupo de sujeitos afetados. Enquanto determinar o grupo de sujeitos diabéticos, obesos ou hipertensos, por exemplo, passa por uma medida de avaliação objetiva, razoavelmente sensível e específica, a determinação de quais são os sujeitos depressivos, desatentos ou com déficit de linguagem requer, freqüentemente, além da exclusão de outros diagnósticos confundidores, o uso de uma série de instrumentos de avaliação, os quais muitas vezes geram resultados dependentes da habilidade do pesquisador.

Frente a esse quadro, é importante analisar alguns dos motivos que tornam tão árdua a tarefa de identificar genes associados a comportamentos ou qualquer outra forma de herança complexa. A principal justificativa é, sem dúvida, a própria interação gênica. No cenário de uma herança monogênica, a mutação de um único gene determina o fenótipo na sua integralidade. Assim, por exemplo, uma mutação no gene que dá origem à proteína distrofina, situado no cromossomo X, provoca a distrofia muscular de Duchenne, doença intratável que leva a fraqueza e atrofia progressiva dos músculos, inclusive do músculo cardíaco. Já nos quadros de herança complexa, o que se observa é uma interação entre dezenas, por vezes até

centenas, de diferentes genes que, é importante notar, não estabelecem uma relação de causalidade direta com a doença. São considerados genes de susceptibilidade ou de predisposição; são genes que participam, mas não determinam, sozinhos, o fenótipo. Além da dificuldade de descobrir quais são esses genes e de identificar que papéis eles desempenham, é preciso ainda saber como esses genes interagem entre si.

Mais do que isso, há que se identificar a presença de componentes ambientais que possam estar participando da manifestação inicial e da determinação subsequente do fenótipo e considerar ainda o modo de ação desses fatores ambientais sobre o genótipo, uma vez que estes podem estar agindo como gatilhos iniciais ou como moduladores, reguladores ou até mesmo inibidores do fenótipo. No caso específico da linguagem, a participação do ambiente – os dados lingüísticos primários aos quais a criança é exposta desde o nascimento – é indubitavelmente decisiva, como comprova o fato de que uma criança filha de pais brasileiros que nascer e crescer no Japão falará japonês (com um domínio da língua que seus pais, por maior que se esforcem, jamais atingirão).

No que diz respeito especificamente às doenças de herança complexa, um ponto importante é a definição dos alelos envolvidos na manifestação do fenótipo e a participação relativa de cada um. Um mesmo gene pode apresentar pequenas variações entre diferentes indivíduos (trocas na seqüência de nucleotídeos), sem que isso tenha efeito danoso. Chama-se alelo cada uma das formas específicas que um gene pode assumir. Genes que apresentam variações alélicas na população com freqüência superior a 1% são denominados polimórficos. O tipo mais comum de polimorfismo é a substituição de apenas um nucleotídeo, troca esta conhecida como SNP (do inglês, *single nucleotide polymorphism*). O estudo desses pontos de variação tem sido responsável pela grande maioria das descobertas sobre as bases moleculares do comportamento.

Postula-se que as doenças complexas obedecem a um modelo de limiar (Fig. 10): cada indivíduo acumularia (herdaria) um número diferente de alelos de susceptibilidade, que, em conjunto com fatores ambientais, determinariam o risco de desenvolver a doença. A soma de fatores genéticos e ambientais, assim, deslocaria o indivíduo para a esquerda (menor risco) ou direita (maior risco) da curva de distribuição normal; abaixo do limiar, espera-se que o indivíduo tenha fenótipo normal e, acima dele, que seja afetado pela doença.

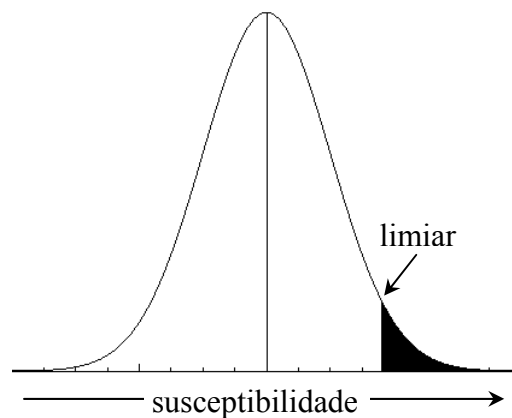


Figura 10. Modelo de limiar para heranças complexas ou multifatoriais. Segundo este modelo, existiria um limiar de fatores genéticos e ambientais para cada condição que, quando ultrapassado, determinaria a manifestação da doença. Quanto mais à esquerda da curva normal, mais “protegido” estaria o indivíduo, já que acumularia um número menor de fatores de susceptibilidade.

Cada indivíduo herda uma combinação única de genes e, a partir do conjunto de genes herdados e dos fatores ambientais observados no meio em que vive, seria possível estimar o risco relativo para que venha a ultrapassar o limiar e desenvolver a doença ou expressar o traço. A identificação de que uma pessoa é portadora de um ou alguns alelos de risco para uma dada doença não significa que ela irá expressar o fenótipo: o acúmulo de muitos outros alelos e a exposição a um estímulo ambiental podem ser necessários para que isso ocorra. Considere, por exemplo, uma dada doença na qual estejam envolvidos 20 alelos. Supondo ainda, hipoteticamente, que cada alelo contribua com uma mesma proporção na determinação do fenótipo, i.e., que cada um tenha o mesmo “peso” na manifestação da doença, pode-se estimar que cada alelo irá contribuir com apenas 5% do fenótipo. Deste modo, dentre os 20

alelos associados à doença, a herança de um único alelo de susceptibilidade pode ser necessária, mas provavelmente não será suficiente para que o indivíduo manifeste a fenótipo.

Desse modo, existem muitos alelos que são necessários para a expressão do fenótipo, mas não suficientes. Os alelos de risco identificados até hoje, em sua maioria, contudo, não são nem necessários nem suficientes. Do ponto de vista metodológico, a identificação de alelos de risco em grupos de indivíduos não-afetados traz dificuldades significativas no que tange à identificação dos genes associados às doenças complexas, já que esses alelos de risco terminam por ser encontrados em ambos os grupos investigados, afetados e normais.

Um terceiro elemento complicador dos estudos para identificação de genes associados a heranças complexas já foi mencionado acima: trata-se da própria dificuldade na definição do fenótipo. Novamente, em oposição ao que se observa nas heranças monogênicas, nas quais o fenótipo é tipicamente bem definido e segue um padrão de tudo-ou-nada, nas heranças complexas há uma ampla variação fenotípica, resultante da distribuição normal na população das diversas variáveis que entram na composição desse fenótipo.

A dificuldade em definir o fenótipo no caso dos déficits lingüísticos é paradigmática e passa pela própria controvérsia na definição do conceito de linguagem. A variabilidade que cerca a definição do fenótipo, de acordo com o entendimento dos diferentes grupos de investigadores, termina por fazer com que estudos em genética do comportamento freqüentemente esbarrem na dificuldade de estabelecer parâmetros uniformes na definição dos grupos de casos e de controles, o que prejudica as possibilidades de replicação dos estudos.

Os distúrbios de linguagem constituem um exemplo bastante representativo dessa dificuldade em estabelecer grupos uniformes com base no fenótipo. O diagnóstico do distúrbio específico de linguagem (SLI, do inglês, *specific language impairment*) é feito por exclusão, ou seja, tem SLI a criança que apresenta problemas no ritmo ou na seqüência de desenvolvimento da fala ou da linguagem que não possam ser explicados por uma outra

condição clínica associada, como problemas auditivos, fenda palatina, retardo mental, distúrbios neurológicos ou qualquer outro transtorno do desenvolvimento (p.ex., autismo). Hoje considerado um distúrbio comum, com prevalência na população estimada entre 6% e 7% (Tomblin et al., 1997), sob o rótulo “SLI” está uma ampla variedade de fenótipos, com natureza e grau de comprometimento lingüístico variável: crianças com dificuldade na aquisição de habilidades expressivas da linguagem, seja essa dificuldade de natureza fonológica, articulatória ou outra; crianças com déficits na compreensão da linguagem (em geral acompanhado de outras dificuldades cognitivas); ou associação de ambos os problemas. Desse quadro resultou a ampla subclassificação atual dos tipos de SLI: distúrbio fonológico, distúrbio de expressão da linguagem e distúrbio misto de expressão-compreensão (Tomblin et al., 1996).

A observação de uma significativa agregação familiar entre os portadores de SLI foi um dos primeiros impulsos para a procura por genes relacionados a distúrbios de linguagem. Comparados à população em geral, membros da família de crianças com SLI têm uma probabilidade entre quatro e oito vezes maior de também serem portadores do transtorno (Tomblin & Pandich, 1999). Recentemente, Tallal et al. (2001), analisando uma amostra de crianças com SLI, encontraram a presença do déficit em 13% das crianças cujos pais eram ambos não-afetados; em 40% das crianças com um dos pais afetado; e em 71,4% das crianças cujos pais eram ambos também portadores de SLI. Embora diversos estudos evidenciem que o SLI tenda a recorrer em membros de uma mesma família, os resultados expressivos dos estudos de agregação podem ser, em parte, atribuíveis a variáveis confundidoras, como fatores ambientais compartilhados pelos membros da família.

A questão de o ambiente estar ou não agindo como variável confundidora na determinação de uma possível etiologia genética pode ser em boa medida esclarecida investigando-se pares de gêmeos, metodologia que se constitui como uma das principais

fontes de evidência da participação de fatores genéticos nos distúrbios de linguagem, como o SLI. Essa estratégia compara a concordância para um determinado traço fenotípico entre gêmeos monozigóticos (MZ) em relação à concordância observada entre gêmeos dizigóticos (DZ). Esse tipo de estudo parte da premissa de que gêmeos MZ e DZ sofrem influência ambiental semelhante; no entanto, os MZ são geneticamente idênticos, enquanto os DZ compartilham apenas a metade de sua informação genética (têm presumivelmente 50% dos genes em comum, como dois irmãos quaisquer). Desse modo, para traços determinados por fatores ambientais, a concordância entre MZ e DZ seria próxima, ao passo que para traços de base fortemente genética a concordância entre MZ seria significativamente maior do que entre DZ. Os resultados desse tipo de estudo genético-epidemiológico permitem, além de demonstrar a existência de um componente genético no transtorno estudado, estimar a herdabilidade (h^2), ou seja, o quanto da variância de um traço ou transtorno é devido a fatores genéticos.

Tomblin & Buckwalter (1998) analisaram 120 pares de gêmeos com “*poor language achievement*” e sem outros transtornos. Como critério de inclusão foi definido que um dos gêmeos deveria apresentar resultados em um teste de linguagem (incluindo tanto habilidades expressivas quanto receptivas), no mínimo, um desvio-padrão abaixo do esperado para sua idade e QI não-verbal. Os resultados mostraram uma concordância de resultados nos testes de 96% entre MZ contra 69% entre DZ. A herdabilidade – quanto da variância nos escores dos testes pode ser atribuída a fatores genéticos – foi estimada em 45%.

Bishop et al. (1995) realizaram estudo semelhante, com 90 pares de gêmeos, utilizando como critério de inclusão uma discrepância ≥ 20 pontos entre os escores dos testes de habilidades verbais e não-verbais. Novamente, encontraram evidências de que fatores genéticos estão envolvidos na etiologia do SLI, ao demonstrar uma concordância entre MZ de 70% contra 46% entre gêmeos DZ.

Desse modo, embora o *FOXP2* permaneça sendo o único gene até hoje descrito como estando diretamente relacionado à linguagem, isso não significa que esta seja a única evidência da participação de genes no desenvolvimento das habilidades lingüísticas. Stromswold (2001), em extensa revisão sistemática da literatura, incluindo mais de 100 estudos com diferentes metodologias (gêmeos, adoção, ligação), observou que fatores genéticos são responsáveis por uma parcela significativa da diferença nas capacidades lingüísticas, tanto entre indivíduos normais quanto entre aqueles com transtorno de linguagem.

Stromswold demonstrou que a concordância entre gêmeos MZ é maior do que entre DZ para transtornos de linguagem escrita e falada. A análise da literatura revelou também que fatores genéticos são responsáveis por mais da metade da variância na capacidade lingüística de indivíduos com transtornos de linguagem. No mesmo sentido, o único estudo de adoção existente indicou que crianças adotadas cujos pais biológicos tinham transtornos de linguagem apresentavam um risco quase três vezes maior de apresentar o mesmo déficit, se comparadas a crianças adotadas cujos pais biológicos não apresentavam déficits lingüísticos. Estudos comparando o desempenho de gêmeos sem transtornos de linguagem em testes lingüísticos mostraram que gêmeos MZ são mais semelhantes entre si do que gêmeos DZ, sendo este resultado independente da habilidade cognitiva geral. Stromswold (2001) encontrou ainda evidências de que fatores genéticos podem influenciar seletivamente habilidades lingüísticas específicas, tais como consciência fonológica, vocabulário e morfossintaxe.

Outra evidência importante do papel de fatores genéticos no desenvolvimento da linguagem vem das diversas síndromes genéticas (das quais a síndrome de Down é a mais conhecida), nas quais a aquisição das habilidades lingüísticas mostra-se prejudicada em diferentes graus e de diferentes formas. Crianças com a síndrome de Down, por exemplo, que resulta de um genótipo com três cromossomos 21 (trissomia do 21), apresentam déficits

articulatórios proeminentes e dificuldades mais marcadas na expressão do que na compreensão da linguagem, com déficits curiosamente semelhantes aos encontrados no SLI (Laws & Bishop, 2004). Déficits lingüísticos variados são também conhecidos nas síndromes de Williams, Angelman, Prader-Willi e X-Frágil, as quais cursam com um nível variável de retardo mental, oscilando desde uma leve dificuldade de aprendizagem até um profundo comprometimento das funções intelectivas.

Em suma, independentemente da identificação precisa de genes, existe um amplo conjunto de evidências sobre a participação de fatores genéticos na etiologia dos transtornos de linguagem e também, em menor escala, em relação à linguagem propriamente dita, em condições não-patológicas. Além da dificuldade em isolar genes especificamente relacionados a funções cognitivas, o conhecimento da estrutura dos genes e das proteínas por eles codificadas pouco pode esclarecer sobre o intrincado funcionamento de funções complexas como a linguagem, que emergem a partir de obscuras interações entre genes e ambiente.

Genes codificam proteínas, não comportamentos. A idéia de que existam “genes da linguagem” ou “genes da gramática” é simplista e pouco realista, se não tola. O que não significa que os genes não tenham um papel fundamental no desenvolvimento dessas habilidades. Múltiplos genes parecem estar envolvidos na organização de estruturas neurais complexas que antecedem e tornam possível o desenvolvimento de habilidades cognitivas complexas, como a linguagem. Assim, sai de cena a questão sobre se os genes participam ou não do processo de aquisição; resta saber como e em que medida eles organizam esse processo. As evidências apresentadas no Capítulo III devem ajudar a esclarecer essas questões.

CAPÍTULO III

ENFIM UM GENE PARA A LINGUAGEM. E DAÍ?

A história de uma descoberta

Uma família, um gene, uma descoberta. Em novembro de 2001, um editorial de uma das mais importantes revistas na área de neurociências propagava a descoberta do gene *FOXP2* como “o primeiro grande triunfo da genética cognitiva” (*Nature Neuroscience*, 2001). Finalmente, o estudo de um traço complexo do comportamento humano – talvez o mais importante deles, a linguagem – havia resultado na identificação precisa de um gene.

O alarde em torno da descoberta do *FOXP2* é justificável. A complexidade de fatores genéticos e ambientais envolvidos na emergência de comportamentos humanos e, principalmente, a intrincada interação entre esses fatores está na base das dificuldades em identificar uma associação específica entre um gene e um comportamento. O anúncio da descoberta do gene *FOXP2*, em 2001, foi precedido por uma longa jornada investigativa que tem início em 1990, quando foi relatado pela primeira vez o caso singular de uma família britânica na qual aproximadamente a metade dos indivíduos das três gerações era afetada por um grave distúrbio de fala e linguagem.

Essa família havia sido identificada pela primeira vez em 1987, quando o diretor da escola onde estudavam alguns de seus membros mais jovens encaminhou-os para especialistas em aconselhamento genético. De imediato, a observação dos 16 membros evidenciou uma dificuldade em organizar e coordenar os movimentos necessários à produção da fala, o que levou Hurst et al. (1990) a estabelecer um diagnóstico de “dispraxia verbal do desenvolvimento”.

A simples recorrência de um distúrbio de linguagem numa mesma família não é, por si só, novidade do ponto de vista científico, uma vez que há ampla evidência de que os casos de SLI tendem a mostrar agregação familiar, como discutido anteriormente. No entanto, na família relatada por Hurst et al., a análise da distribuição de indivíduos afetados e não-afetados revelou um quadro raro e inesperado de herança para um distúrbio lingüístico. Diferentemente dos muitos relatos prévios da literatura, a composição do heredograma (Fig. 11) da família KE – como ficou conhecida – era compatível com um padrão de herança mendeliano dominante e autossômico, uma vez que todos os indivíduos afetados tinham ao menos um progenitor afetado, estando o fenótipo presente nas três gerações (características típicas de heranças dominantes), distribuindo-se igualmente entre homens e mulheres (característica típica de herança autossômica, i.e., não-relacionada aos cromossomos sexuais).

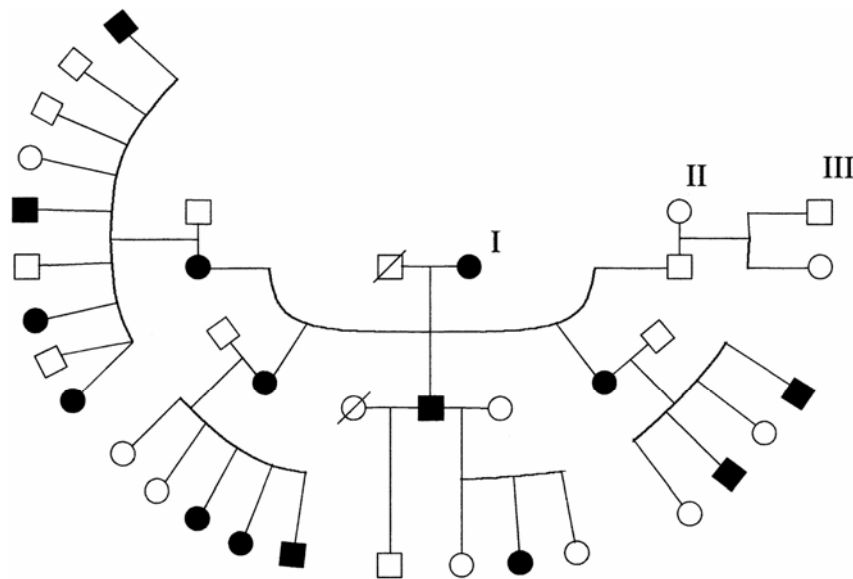


Figura 11. Heredograma da família KE. Círculos representam mulheres; quadrados representam homens. As figuras sombreadas indicam indivíduos afetados e as não-sombradas, indivíduos não-afetados. As barras sobre as figuras indicam indivíduo já falecido (extraído de Watkins et al., 2002).

Tal padrão de herança sugeriu aos pesquisadores a possibilidade de se tratar de uma herança determinada pela mutação de um único gene, originalmente encontrado na avó (geração I), transmitido a quatro dos cinco filhos (geração II) e repassado a aproximadamente 50% dos netos (geração III).

O aparente caráter monogênico do transtorno de linguagem apresentado pela família KE levou um grupo britânico de pesquisadores – liderados por Simon Fisher e sediados no prestigioso *Wellcome Trust Center for Human Genetics* – a investir na possibilidade de localizar o gene (e a mutação) responsável por tal manifestação. Conforme discutido anteriormente, a identificação de genes envolvidos em distúrbios monogênicos é, se comparada à pesquisa de distúrbios que derivam de heranças complexas, mais simples, mais direta e mais promissora.

A própria revelação da existência da família KE, antes mesmo das sofisticadas análises que estavam por vir para definir o fenótipo e esclarecer o genótipo, traz duas evidências preliminares importantes. A primeira é a evidência da presença de um componente hereditário na etiologia de um transtorno de linguagem. Isso porque o padrão específico de agregação familiar, neste caso, é muito significativo e dificilmente atribuível a fatores ambientais. Considerando que esses indivíduos compartilham o mesmo ambiente (visitam-se uns aos outros ou vivem na mesma casa, moram na mesma cidade, freqüentam as mesmas escolas, têm amigos em comum), como explicar que uma metade desenvolve a linguagem normalmente e a outra não? Outra evidência preliminar importante materializa um antigo postulado da lingüística gerativa de que o fator ambiental o qual se presume interagir com os genes envolvidos no processo de desenvolvimento da linguagem – i.e., a exposição a dados lingüísticos primários – é necessário, mas não é suficiente para a determinação do fenótipo, uma vez que os indivíduos afetados e não-afetados da família KE foram expostos a um ambiente lingüístico comum nos primeiros anos de vida.

Na busca pelo gene mutado em metade dos indivíduos da família KE, os pesquisadores empregaram uma série de marcadores – pequenas seqüências de DNA de localização conhecida no genoma humano – a fim de identificar uma correlação entre a segregação desses marcadores nos indivíduos afetados e nos não-afetados e a presença ou a

ausência do distúrbio. Assim, se um determinado marcador genético for herdado por todos os membros afetados da família, mas não pelos não afetados, é provável que o gene responsável pela doença esteja localizado nas proximidades desse marcador.

A estratégia empregada pelo grupo britânico baseia-se no conceito de ligação genética (*linkage*) – propriedade de dois ou mais *loci* (posição ou lugar ocupado pelo gene) em um mesmo cromossomo que, por proximidade física, tendem a ser mantidos unidos em um mesmo gameta e, por isso, tendem a ser herdados juntos. Embora todo ser humano herde duas cópias de cada cromossomo, uma do pai e uma da mãe, de geração em geração, parte dessas cópias vai sendo misturada e modificada, num processo aleatório conhecido como recombinação gênica, que ocorre principalmente nos estágios iniciais do desenvolvimento do zigoto. Assim, o genoma de um indivíduo não contém cromossomos exatamente iguais ao do pai ou da mãe, mas semelhantes, podendo ser caracterizados como uma “colcha de retalhos” dos genes das gerações precedentes. Essa variabilidade torna-se útil para a identificação de genes, uma vez que permite “fragmentar” o cromossomo em regiões que são comuns (ou não) aos membros de uma dada família.

Um marcador identificado como estando consistentemente associado aos indivíduos afetados não estabelece, obrigatoriamente, uma relação de causalidade direta com o distúrbio investigado, ou seja, não necessariamente a seqüência de DNA identificada nos afetados é parte da seqüência mutada. A expectativa é de que a seqüência mutada esteja, pelo menos, nas proximidades do marcador, de modo que este funcione como um índice para a localização cromossômica do gene mutado. Em 1998, Fisher e colaboradores anunciaram a identificação de um marcador que co-segregava perfeitamente com o distúrbio, localizado no cromossomo 7 (7q31) e deram à região o nome de SPCH1 (Fisher et al., 1998) (Fig. 12).

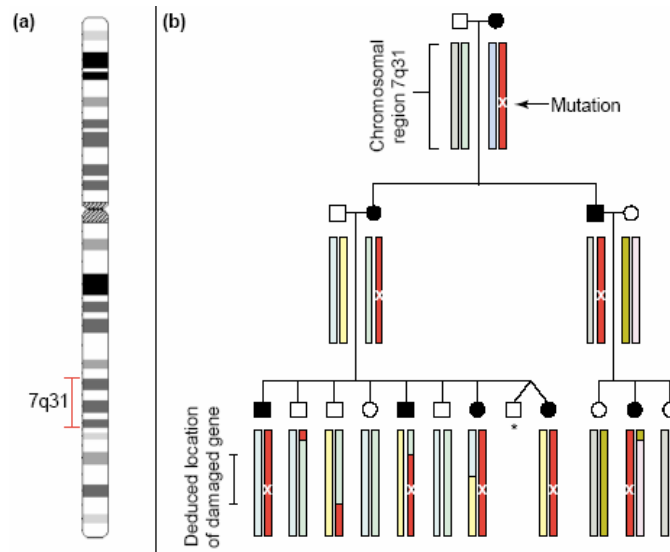


Figura 12. A localização do gene mutado na família KE. Para simplificar, apenas parte da família está representada. O asterisco indica um gêmeo não-identico cujo DNA não esteve disponível para o estudo. As diferentes cores indicam as diferentes origens das cópias dos cromossomos na família KE. Descobriu-se que todos os indivíduos afetados haviam herdado uma cópia idêntica de uma porção do braço longo do cromossomo 7 (7q31). Essa cópia (indicada pela cor vermelha) teve origem em um cromossomo da avó materna (extraída de Marcus & Fisher, 2003).

Apesar da importante descoberta, a localização do gene especificamente mutado na família KE ainda poderia levar anos. Isso porque, segundo estimativa dos próprios pesquisadores, o intervalo delimitado pela região SPCH1 (*speech 1*) continha entre 50 e 100 genes. Restava, então, iniciar a investigação dos genes, um a um, começando por aqueles que, através de outros estudos, já haviam sido relacionados à atividade cerebral. O longo caminho até a identificação precisa do gene responsável pelo distúrbio observado na família KE foi subitamente reduzido, em 2000, quando um golpe de sorte facilitou largamente o trabalho dos pesquisadores.

Paralelamente à investigação dos genes localizados na região SPCH1, os pesquisadores britânicos iniciaram uma busca por indivíduos com distúrbios de linguagem que apresentassem algum tipo de anomalia cromossômica envolvendo o cromossomo 7. Isso porque, historicamente, a identificação de grandes alterações estruturais, como as observadas em nível de cromossomos, tem auxiliado no mapeamento de distúrbios monogênicos. Um paciente, conhecido como CS, foi encaminhado ao grupo de Fisher apresentando uma grande

dificuldade na articulação dos sons da fala, acompanhada por um significativo atraso no desenvolvimento de habilidades de expressão e compreensão da linguagem.

À época com cinco anos e meio, CS havia sido diagnosticado previamente ao seu nascimento, via amniocentese, como sendo portador de uma translocação cromossômica envolvendo os braços longos dos cromossomos 5 e 7 – $t(5;7)(q22;q31.2)$. Ou seja, esses cromossomos haviam se “quebrado” em duas partes e se rearranjado, de modo que parte do material genético do cromossomo 5 estava no cromossomo 7 e vice-versa. Ao nascimento, CS não apresentou qualquer anormalidade, mas por volta dos dois anos de idade voltou a ser encaminhado aos médicos em função de um atraso no desenvolvimento da linguagem e de um leve retardo no desenvolvimento motor (Lai et al., 2000). A análise genética da localização exata do ponto de ruptura do cromossomo 7 em CS revelou que esta – fortuitamente – envolvia a região SPCH1.

Percebendo a feliz e rara convergência do caso CS com a família KE, os pesquisadores voltaram-se, então, para a investigação do ponto específico, dentro da região SPCH1, afetado pela translocação no menino CS. Lai et al. (2000) relataram que a ruptura no cromossomo 7 de CS ocorria muito próxima a um gene parcialmente mapeado (CAGH44), cuja proteína era expressa no tecido cerebral. No ano seguinte, o mesmo grupo de pesquisadores finalizou o sequenciamento desse gene.

A análise da seqüência completa do gene e a comparação desta ao genoma da família KE revelaram que todos os membros afetados apresentavam no seu DNA a troca de um único par de bases, um SNP, com a substituição de G por A, ou seja, de uma guanina por uma adenina. Essa substituição estava localizada em uma região codificadora do gene (éxon 14) e resultava na troca de um aminoácido na seqüência da proteína (substituição de uma arginina por uma histidina). Como a presença dessa arginina é invariável nos demais genes da família *FOX*, ela provavelmente exerce uma função importante, sugerindo que a sua substituição

resulte em perda de função da proteína codificada. Além de co-segregar perfeitamente com a presença do distúrbio de linguagem nos indivíduos da família KE, a troca de uma guanina por uma adenina, no éxon 14 do gene, não estava presente em nenhum dos 364 cromossomos de indivíduos normais investigados como controles. Em função das características estruturais da proteína por ele codificada, identificou-se o gene como pertencente à ampla família de genes *FOX* e, de acordo com as regras de nomenclatura, foi chamado *FOXP2*.

Como discutido no Capítulo II, o *FOXP2* é um gene que codifica uma proteína pertencente à classe dos fatores de transcrição. Trata-se, portanto, de proteínas que se ligam a e interagem com as regiões regulatórias de outros genes, modulando seus níveis de transcrição e determinando, em última análise, a abundância e a disponibilidade das proteínas formadas por esses genes. Os fatores de transcrição são comumente comparados a interruptores ou chaves de luz: eles ligam ou desligam outros genes que estão “em rede” com ele. Na verdade, sua função pode ser melhor comparada com a de um termostato, já que não se trata de uma ativação tudo-ou-nada, mas da regulação ou da modulação de níveis de transcrição (Ridley, 2003).

Desde a descoberta original do gene em embriões de mosca-das-frutas, genes pertencentes à família *FOX* foram descobertos em diversas espécies, incluindo seres humanos, mas só recentemente o seu papel em variados processos do desenvolvimento vem sendo compreendido. Os genes da família *FOX* foram agrupados em função de as proteínas que codificam apresentarem uma estrutura tridimensional característica em seu domínio de ligação ao DNA. Essa região através da qual a proteína se liga ao DNA é de importância crucial, pois é essa ligação que permite que as proteínas *FOX* regulem a expressão de outros genes; no caso do *FOXP2*, a mutação afeta justamente esse domínio de ligação ao DNA. A família de genes *FOX* está subdividida em 17 classes (de A a Q). Esses genes estão presentes tanto em animais vertebrados quanto invertebrados, mas o número de genes parece ter

aumento ao longo da evolução (Lehmann et al., 2003). Essa família de genes caracteriza-se por uma ampla variedade funcional, desempenhando diversos papéis reguladores da diferenciação e proliferação celular e da transdução de sinais químicos, com diversos membros atuando como importantes reguladores da embriogênese. Muitos já foram relacionados a doenças do desenvolvimento humano, incluindo glaucoma (*FOXC1*), agenesia da tireóide (*FOXE1*) e em diferentes tipos de imunodeficiência (*FOXN1* e *FOXP3*) (Carlsson & Mahlapuu, 2002).

No que tange à proteína *FOXP2* em particular, sabe-se que sua expressão não está restrita ao tecido cerebral. Embora seja expressa em altos níveis em regiões circunscritas do cérebro durante o desenvolvimento (como discutido a seguir) – o que, especulativamente, deve estar relacionado à formação das estruturas neurais necessárias ao processo de aquisição da linguagem – a proteína *FOXP2* também é expressa no coração, nos pulmões, no intestino e em outros órgãos no indivíduo adulto (Lai et al., 2001).

Tal achado não deve causar surpresa, uma vez que se trata de um fator de transcrição – e, como tal, interage com outros tantos genes importantes no funcionamento de coração, pulmão, intestino, etc. É legítimo, no entanto, questionar por que os problemas apresentados pela família KE restringem-se apenas a linguagem, se essa proteína é também importante para o desenvolvimento de outros órgãos vitais. A resposta, segundo Lai et al. (2001), seria uma questão de dosagem: todos os indivíduos afetados da família britânica herdaram uma cópia mutada do gene, sendo a outra absolutamente normal. Assim, sugere-se que, para certos tecidos, a quantidade de proteína expressa pelo único gene intacto seria suficiente para garantir seu funcionamento adequado, enquanto para um desenvolvimento cerebral normal, esse nível crítico de expressão não poderia ser alcançado com um único gene ativo. Essa hipótese é reforçada pela observação de altos níveis de expressão do *FOXP2* no cérebro de fetos humanos com desenvolvimento normal (Lai et al., 2003).

Nas próximas seções, são descritos os achados relativos ao fenótipo dos indivíduos afetados da família KE, com o objetivo de melhor compreender o papel do *FOXP2* no desenvolvimento da linguagem. Os achados estão divididos em duas categorias, destacando a caracterização dos déficits de linguagem e as alterações neuroanatômicas e neurofisiológicas associadas à mutação desse gene. As diversas descrições da família publicadas desde 1990 revelaram significativas inconsistências entre as análises efetuadas, o que terminou por tumultuar a caracterização fenotípica dos KE (Fisher et al., 2003). Apesar disso, os dados revelam um panorama rico e complexo de fenômenos que emergem a partir da mutação, contribuindo principalmente para a discussão das relações de (in)dependência que se estabelecem entre linguagem e outros domínios cognitivos humanos.

A definição do déficit nuclear como centro da discórdia

O déficit articulatorio – traço mais evidente no fenótipo dos membros afetados da família – foi originalmente relatado por Hurst et al. (1990) como uma forma de apraxia verbal do desenvolvimento. Os autores observaram seis membros da família, com idades à época entre oito e dezoito anos, que foram assim descritos como conjunto: *“the children had serious communication difficulties and their speech disorder was classified as a severe form of developmental verbal apraxia, since both speech and expressive language were involved”* (Hurst et al., 1990, p.352). O termo apraxia origina-se do grego (falta de ação) e sugere uma perturbação na habilidade de transmitir ou expressar uma resposta motora complexa em uma modalidade específica. No caso específico da fala, trata-se de um distúrbio sensoriomotor lingüístico, de articulação, caracterizado por prejuízo na capacidade de programar a posição adequada dos músculos da fala e o seqüenciamento dos demais movimentos musculares necessários à produção de fonemas (Nicolosi et al., 1996).

Tal definição está em acordo com a descrição apresentada por Hurst et al., que relata a dificuldade das crianças com a articulação de sons complexos, como encontros consonantais (p.ex., ‘boon’ para *spoon* e ‘bu’ para *blue*), e a tendência a omitir fonemas no início de palavras (p.ex., ‘able’ para *table*). Há também a descrição de uma tendência à omissão de “sons” em final de palavras (possivelmente morfemas), assim como a tornar palavras polissilábicas em di- ou monossilábicas. Os enunciados são curtos, incluindo a combinação de não mais do que três palavras por volta dos cinco, seis anos de idade. À exceção da linguagem, outros marcos do desenvolvimento infantil (como sentar e caminhar) foram observados razoavelmente dentro da idade esperada. Todas as crianças analisadas apresentaram atraso no desenvolvimento da linguagem, com emissão das primeiras palavras por volta dos três anos.

Nessa primeira descrição, fica claro que a dificuldade com a linguagem na família KE não se limita a problemas articulatórios. Assim, Hurst et al. observam uma demora significativa em atividades de nomeação de objetos comuns (acesso lexical), acompanhada por uma tendência ao “uso de palavras aproximadas”, como *glass* ou *tea* para ‘*cup*’ e *sky* para ‘*star*’. Os autores relatam ainda a não-espontaneidade da fala, com as crianças mostrando-se relutantes em conversar (sic). Embora tenham enfatizado os déficits expressivos, os autores observaram também dificuldades na compreensão da linguagem, notadamente com frases comparativas e passivas: “‘*The knife is longer than the pencil*’ was poorly understood. ‘*The girl is chased by the horse*’ became ‘*the girl is chasing the horse*’, and ‘*the boy chasing the horse is fat*’ was interpreted as being the same as ‘*a thin boy chasing a fat horse*’” (p.354).

Embora a metodologia empregada por essa equipe de pesquisadores para avaliar a família não esteja clara, os dados apresentados sugerem, de fato, pelo menos, a presença de uma dispraxia, com efeito marcante sobre a linguagem, como pode se depreender dos exemplos acima. É importante observar que Hurst et al. classificaram a dispraxia como verbal

– e não oral – implicando que a dificuldade estaria restrita à articulação de sons da fala. Contudo, tal classificação esbarra em observações relatadas pelos próprios autores de que os indivíduos analisados eram capazes de posicionar a língua e os lábios para executar movimentos simples, mas eram incapazes de realizar seqüências complexas de movimentos orofaciais conforme o comando dado pelo entrevistador. Esse achado sugere uma dificuldade na coordenação e no seqüenciamento de movimentos orofaciais – não somente aqueles necessários à fala, mas uma dificuldade mais geral com o concatenamento desses movimentos.

Atentos a esse achado, Vargha-Khadem et al. (1995) levantaram a possibilidade de que essa dificuldade na repetição de movimentos orofaciais complexos fosse ocasionada não pela dispraxia, mas por uma dificuldade na própria compreensão do comando verbal. Ao invés de instruir verbalmente os indivíduos afetados sobre quais movimentos deveriam ser repetidos, os pesquisadores realizaram um teste de imitação. Os resultados refutaram a hipótese de incompreensão do comando verbal, uma vez que ficou demonstrado que os indivíduos afetados da família KE eram capazes de imitar movimentos orofaciais simples (e, portanto, de compreender o comando verbal “imite”), mas demonstravam grande dificuldade com movimentos simultâneos ou seqüenciais rápidos, como, por exemplo, fechar os lábios, abrir a boca e pôr a língua para fora.

No ano seguinte à publicação original do grupo de Hurst, Gopnik & Crago (1991) publicaram a primeira análise propriamente lingüística dos membros da família. As autoras foram também as primeiras a argumentar que a perturbação extrapolava o domínio da fala e da articulação dos sons, afetando também – e principalmente – a esfera gramatical. Coerentemente com essa posição, rejeitaram a classificação de apraxia, optando em seu lugar pelo termo disfasia, que denomina um distúrbio lingüístico causado por dano cerebral e caracterizado por prejuízo parcial da compreensão e/ou uso da linguagem, estando ausentes

déficits sensoriais primários, deterioração mental geral ou distúrbios psiquiátricos (Nicolosi et al., 1996).

A análise de Gopnik & Crago contou com a participação de 20 membros das três gerações da família KE (13 disfásicos), com idades variando entre 2 e 74 anos, idade do mais antigo membro afetado da família – a avó –, hoje falecida. Com exceção desta, todos os demais indivíduos haviam sido encaminhados na infância, pela escola, para acompanhamento fonoaudiológico, sendo submetidos a tratamento durante um ou mais anos. Todos os indivíduos afetados da terceira geração (netos) freqüentaram, desde os quatro anos de idade, uma escola especializada em crianças com atraso no desenvolvimento da linguagem. Segundo o relato da avó, os seus filhos afetados apresentavam na infância problemas de fala bastante semelhantes entre si e semelhantes àqueles agora observados entre os netos; também os pais referiram esse quadro, afirmando que seus filhos afetados mais velhos, quando jovens, apresentavam uma fala semelhante a dos caçulas disfásicos – todos referindo uma tendência à melhoria do quadro a partir dos sete anos de idade aproximadamente.

Essas observações mostram que a trajetória no desenvolvimento da linguagem dos membros afetados da família KE segue um curso, ao menos superficialmente, homogêneo, com atenuação dos sintomas no início da vida escolar (este fato está também de acordo com a avaliação de Hurst et al). Como todos, exceto a avó, receberam alguma forma de terapia fonoaudiológica após a entrada na escola, é possível que essa melhora seja decorrente do treinamento específico, não estando relacionada a algum aspecto próprio do desenvolvimento. O grau de melhora do quadro, contudo, é incerto: segundo Gopnik & Crago (1991), à idade adulta, o déficit seria “praticamente imperceptível”, sendo revelado apenas através de testes lingüísticos específicos; por outro lado, os pesquisadores são unânimes ao afirmar que, na seleção de casos e controles, jamais houve discordância sobre quais membros eram afetados e

quais não eram afetados pelo distúrbio, de modo que as dificuldades articulatórias parecem permanecer razoavelmente evidentes.

Gopnik e Crago (1991) utilizaram testes adaptados, originalmente desenvolvidos para avaliação da linguagem em indivíduos com afasia adquirida. Os testes foram aplicados individualmente, em 1989, na casa dos entrevistados e todo o material foi registrado em fitas de áudio. Os sete membros não-afetados da família que participaram como controles no estudo eram da terceira geração. No entanto, dos 20 indivíduos analisados, apenas os resultados para seis casos e seis controles são apresentados. A justificativa das autoras é a de que, ao selecionar apenas os indivíduos mais velhos da amostra, obtém-se um quadro mais preciso do estado final da gramática desses indivíduos, admitindo que todos já atingiram um estado final de competência gramatical. Assim, a amostra final é composta por indivíduos com as seguintes idades:

- afetados: 16, 17, 40, 42, 45 e 74 (idade média: 39 anos);
- não-afetados: 8, 12, 13, 14, 15 e 17 (idade média: 13 anos).

Note-se que as autoras alegam ter realizado um pareamento por idade, mas este está prejudicado no estudo, como evidenciam as médias de idade nos dois grupos. Além disso, embora a justificativa para a eliminação dos dados dos mais jovens (primeira geração) seja pertinente, essa exclusão sumária compromete o estudo, uma vez que há evidências de que o déficit articulatório/orofacial – ao qual Gopnik e Crago não dão saliência – é mais evidente nos membros de menor idade.

Na primeira tarefa, os indivíduos da família KE foram testados quanto à sua habilidade de discriminar a categoria de plural (marcada por *-s*). Sobre uma mesa, foram distribuídos aleatoriamente objetos (como livros, giz de cera e balões) em pilhas (plural) ou na forma de um único objeto isolado (singular). Foram testados comando simples (ex.: *“Please, touch the*

books”), que exigiam que apenas um tipo de objeto fosse tocado, e comandos complexos (ex.: “*Put the crayon on the balloons*”), que exigiam a observação da forma de duas palavras quanto à classe singular ou plural. Para os comando simples, a média dos resultados foi idêntica entre os dois grupos, enquanto para os comandos complexos o grupo não-afetado saiu-se ligeiramente melhor que o grupo de disfásicos, mas sem diferença significativa entre os dois grupos ($p=0,36$). Ao demonstrar que os membros afetados da família KE são sensíveis à marcação de plural, i.e., são capazes de perceber a diferença entre as formas marcadas (“-s”) e não-marcadas e associá-las a diferenças de número, as autoras procuram afastar a hipótese de déficit de processamento auditivo como origem das dificuldades de linguagem da família.

Uma pequena digressão é importante para esclarecer a relevância desse achado. Desde que os primeiros dados sobre a família KE vieram a público, o déficit foi considerado pela maioria dos pesquisadores como um subtipo particular de SLI. Como discutido anteriormente, o SLI é um diagnóstico de exclusão, de modo que fenótipos consideravelmente diversos acabam sendo inclusos sob o mesmo rótulo, justificando, em parte, a inclusão dos KE nessa categoria. É interessante notar, inclusive, que recentemente vem-se admitindo uma fusão da chamada dislexia do desenvolvimento (lembrar da classificação de Hurst et al., 1990) com o SLI, antes consideradas duas entidades diagnósticas distintas (Bishop & Snowling, 2004). Embora crianças com SLI possam apresentar déficits fonológicos, morfológicos e/ou sintáticos variados, um dos marcos centrais do SLI é uma dificuldade no uso e na compreensão de artigos, definidos e indefinidos, verbos auxiliares, marcas de passado (“-ed”), terceira pessoa do singular e plural (“-s”). Esses achados foram replicados em diferentes línguas (ver Clahsen et al., 1997; Clahsen & Dalalakis, 1999).

Observando que as crianças com SLI frequentemente omitiam de sua fala espontânea as marcas de plural (“-s”) e passado regular (“-ed”), Leonard (1989) propôs a chamada “*Perceptual Deficit Hypothesis*” segundo a qual esses itens seriam omitidos da fala em função

de sua relativa ausência de substância fonética. A omissão das marcas flexionais decorreria, assim, do fato de que esses morfemas são tipicamente não-salientes de um ponto de vista fonético, independentemente de seu conteúdo gramatical. Estudos subsequentes fortaleceram essa hipótese ao identificar déficits no processamento de informações auditivas nas crianças com SLI (ver, p.ex., Ziegler et al., 2005). Naturalmente, essa hipótese de que o SLI teria como origem um déficit de processamento – não sendo, portanto, um problema exclusivamente lingüístico – não foi aceita por toda a comunidade de pesquisadores. Daí o interesse de Gopnik e Crago (1991) em descartar uma possível dificuldade de processamento auditivo nos KE.

Uma vez atestado que os indivíduos disfásicos da família KE são capazes de reconhecer marcas de plural, Gopnik e Crago partiram para uma avaliação da capacidade de formação do plural, aplicando um teste com pseudopalavras da língua (*nonsense words*). Pseudopalavras são estímulos verbais construídos dentro das regras estruturais de uma língua, i.e., itens que podem ser lidos, escritos e repetidos, mas que não apresentam um conteúdo semântico no léxico atual daquela língua (Baddeley et al., 1998). A razão para o emprego de pseudopalavras, neste teste, é que as formas plurais dessas palavras não correm o risco de ter sido memorizadas (armazenadas no léxico); como essas palavras nunca foram ouvidas antes, a construção de sua forma plural deve, necessariamente, advir da aplicação de algum tipo de regra. O teste consistia em apresentar aos indivíduos da família KE a figura de um animal imaginário, ao lado de uma outra figura com a representação de vários exemplares daquele mesmo animal. Para eliciar a formação do plural, a pesquisadora apontava para a figura e dizia “*This is a zoop*” e, então, apontava para a outra figura, dizendo “*These are ...*”. O procedimento foi repetido com seis pares de figuras, cujas terminações plurais terminavam com consoantes surdas (“*zoop*” e “*zat*”), sonoras (“*tob*” e “*wug*”) ou sibilantes (“*zash*” e “*sas*”).

Dentro do modelo de faculdade de linguagem proposto pela teoria gerativa, discutido no Capítulo I, as línguas se organizam em torno de um sistema computacional, invariável e geneticamente determinado, capaz de manipular símbolos de acordo com princípios previamente estabelecidos, e de um léxico, repositório das idiossincrasias lingüísticas, que devem ser aprendidas e armazenadas. Esse modelo foi desenvolvido principalmente através de estudos extensivos sobre a formação do passado em inglês, construção esta capaz de evidenciar a atividade dos dois subsistemas. Existiriam dois mecanismos de formação do passado: o primeiro, referente à flexão regular, como em *play* → *played*, está baseado em uma regra computacional que concatena “V + *-ed*”, adicionado o sufixo à raiz do verbo; o segundo, que diz respeito à flexão de verbos irregulares, como no exemplo *go* → *went*, depende da recuperação de formas armazenadas individualmente no léxico mental. Assim, esses mecanismos operariam sob o princípio de que, uma vez recuperada uma forma armazenada de passado, a aplicação da regra computacional é bloqueada. Desse modo, a linguagem seria, ao mesmo tempo, um produto de regras computacionais e da memória associativa, estando o funcionamento dos dois subsistemas dissociado um do outro (Pinker 1991, 2002). A literatura sobre aquisição e processamento da linguagem é rica em evidências a favor dessa abordagem, conhecida como *Words and Rules* (para uma revisão ver Marcus et al., 1995).

Os resultados de Gopnik e Crago mostraram uma diferença bastante significativa entre indivíduos afetados e não-afetados ($p=0,0078$) nos testes de formação do plural. A heterogeneidade dos erros cometidos pelos indivíduos disfásicos da família KE nesta tarefa sugere que, apesar da capacidade de distinguir formas com e sem marcação de plural e de compreender o conceito, esses indivíduos possivelmente não tenham uma regra internalizada para a formação de plural. Muito embora as autoras não apresentem todas as respostas dadas, a amostra do Quadro 3 abaixo ilustra tal consideração. Vale notar ainda, a título de

curiosidade, que as autoras relatam que uma adulta testada, quando apresentada ao primeiro item, fez uma longa pausa antes de fornecer a resposta, repetindo o enunciado em voz baixa diversas vezes até sussurrar para si mesma “*add an ‘s’*”.

Pseudopalavras	
<i>singular</i>	<i>plural</i>
sas	sasss
wug	wug
zat	zacko
tob	tobez

Quadro 3. Amostra de erros cometidos pelos indivíduos afetados da família KE em tarefa de formação do plural de pseudopalavras no inglês. Dados de Gopnik & Crago (1991).

A hipótese apresentada no artigo sugere que esta seja uma primeira evidência do que as autoras denominaram “*feature-deficit theory*”. Segundo essa teoria, esses indivíduos apresentam problemas no processamento de traços gramaticais (marcas de pessoa, tempo e número) porque suas gramáticas não se valeriam de regras internalizadas para a análise e decomposição de itens lexicais e, conseqüentemente, não gerariam itens lexicais marcados com tais traços. Os itens lexicais seriam então interpretados como um todo, uma unidade, uma vez que as marcas de flexão não são entendidas como oriundas de uma regra abstrata de generalização:

Under this hypothesis individual “s”-marked forms are, for these dysphasics, simply unanalyzed lexical items that include “numerous” in the specification of their meaning. “Books” therefore is interpreted by them as an unanalyzed lexical item meaning “several reading objects” rather than being generated by the morphological rules from an underlying representation of the form book + plural (Gopnik & Crago, 1991:36).

A capacidade de processamento da linguagem, assim como a compreensão geral de comandos complexos seqüenciais, parece não estar prejudicada entre os indivíduos disfásicos da família KE. Frente a instruções como “*Here are three crayons. Drop the yellow one on the floor, give me the blue one, and pick up the red one*” não foi observada diferença significativa

na resposta dos dois grupos, com quase todos os indivíduos respondendo ao comando com perfeição ($p=0,097$). Esse achado reafirma o predomínio do déficit na esfera de expressão ou recepção da linguagem e sugere que esses indivíduos não apresentam limitações na compreensão por sobrecarga processual.

Também as habilidades de compreensão sintática foram avaliadas por Gopnik e Crago (1991). A partir de ilustrações, foram feitos testes contrastando o uso de pronomes reflexivos e não-reflexivos (“*He washes him*” versus “*He washes himself*”); pronomes pessoais singulares e plurais (“*He holds him*” e “*He holds them*”); orações ativas e passivas na forma negativa (“*The truck does not pull the car*” e “*The truck is not pulled by the car*”); e pronomes possessivos (“*Show me the mother’s baby*” e “*Show me the baby’s mother*”), os quais não mostraram diferenças significativas entre os dois grupos, contrastando com a observação de Hurst e colaboradores.

Além da assimetria entre habilidades expressivas (prejudicadas) e receptivas (preservadas), as autoras evidenciaram ainda uma dificuldade significativa com julgamentos de gramaticalidade (Fig. 13). O comando instruía os indivíduos a informar se a frase dita pelo entrevistador era uma frase correta ou não na língua inglesa. Foram apresentadas 30 frases, das quais 21 continham variados tipos de erros – erros de número, pessoa e tempo (p.ex., “*The boy eats three cookie*” ou “*The boy kiss the pretty girl*”). Além disso, sempre que os sujeitos julgassem a frase como incorreta, era-lhes solicitado que a corrigissem.

A correção das frases julgadas agramaticais também mostrou diferenças significativas entre afetados e controles ($p=0,00001$). A *feature-deficit theory* construída por Gopnik e Crago ganha aqui mais uma informação importante: embora esses indivíduos reconheçam, acusticamente, a presença de traços formais (como evidenciado pelo primeiro teste, que testava distinção de singular/plural) e consigam atribuir-lhe algum sentido (naquele caso, algo como “mais do que um”), a função sintática desses elementos parece não ser reconhecida

pelos membros afetados da família KE. Esse mesmo tipo de dificuldade na compreensão da função de traços formais da língua, como àqueles necessários ao estabelecimento dos padrões de concordância de pessoa e tempo, é tradicionalmente identificada nos indivíduos com SLI, e a explicação para esse déficit é objeto de amplo debate na literatura.

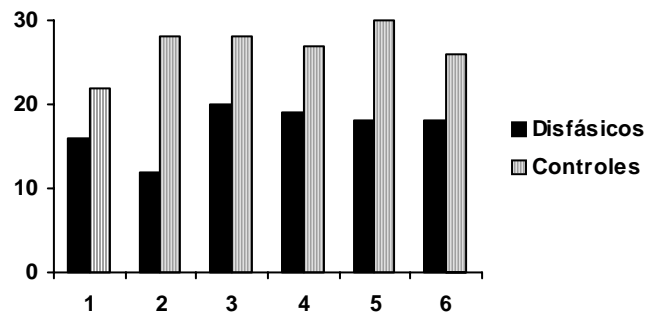


Figura 13. Resultados dos doze indivíduos testados, pareados por idade, no teste de julgamento de gramaticalidade de 30 frases. A análise dos resultados mostrou uma diferença significativa entre os indivíduos afetados e os controles da família KE ($p=0,00001$). Adaptado de Gopnik & Crago (1991).

É interessante notar, contudo, que em frases agramaticais contendo erros na estrutura argumental – avaliadas numa tarefa diferente – não houve diferença significativa entre os dois grupos. Toda frase tem um núcleo que determina o número e os tipos de argumentos (lugares) necessários; assim, por exemplo, na frase “Paulo enviou o livro ao amigo”, o verbo enviar é o núcleo, o qual determina a necessidade de três argumentos:

- Paulo: aquele que envia; papel semântico ‘agente’; função gramatical ‘sujeito’;
- amigo: aquele para quem se envia; papel semântico ‘alvo’; função gramatical ‘objeto indireto’;
- livro: aquilo que é enviado; papel semântico ‘tema’; função gramatical ‘objeto direto’.

De acordo com a teoria de Princípios e Parâmetros (Chomsky, 1981), o conhecimento da estrutura argumental de uma frase, a partir do reconhecimento de seu núcleo, é função do léxico, que não especifica apenas o significado de uma palavra, mas também informações

categoriais (se é um nome, verbo ou adjetivo) e sua seleção semântica (argumentos). Desse modo, esse achado reafirma a proposta das autoras de que o déficit da família KE está seletivamente relacionado às regras (ao sistema computacional), com preservação da memória associativa lexical. Entre as frases agramaticais investigadas estavam as seguintes:

- “*The girl eats a cookie to the boy*”;

- “*The boy puts the book*”;

- “*The nice girl gives*”.

Os indivíduos disfásicos também obtiveram resultados significativamente inferiores aos membros não afetados da família em testes de morfologia derivacional (“*There’s a lot of sun*” – “*It’s very _____*”), sendo que os indivíduos disfásicos mais jovens (com 16 e 17 anos) não produziram a resposta esperada para nenhuma das 10 estruturas apresentadas pelos investigadores. A morfologia derivacional corresponde aos processos pelos quais se torna possível alterar a classe gramatical de uma palavra, através do acréscimo ou remoção de afixos. Respostas semanticamente apropriadas, mas morfologicamente não relacionadas ao item de estímulo não foram aceitas (p.ex., “*You can wash this. It is dirty*”), tendo sido freqüentemente a escolha de resposta do grupo disfásico.

Outra diferença bastante expressiva entre os dois grupos foi verificada na tarefa de flexão verbal (Fig. 14). Exemplos das 10 frases testadas incluem “*Yesterday the girl baked a cake. Tomorrow she _____*” e “*The boy always cries. Right now he _____.*” Além desses verbos, as respostas exigiam a formação de passado simples (*kissed* e *walked*), passado irregular (*went* e *was*), futuro, presente e presente contínuo. A média de acertos, sobre um total de 10, ficou entre os controles em 7,83, contra 3,16 entre os afetados ($p=0,0023$).

Freqüentemente, reforçando a teoria das autoras, as respostas dos indivíduos afetados apresentavam-se como uma forma não-marcada do verbo elicitado, o que não foi observado

entre nenhum dos controles, ressaltando também mais uma vez a semelhança do déficit dos KE com as descrições prévias de indivíduos com SLI. Outro erro observado foi a troca do verbo-estímulo por um verbo semanticamente relacionado, erro também verificado entre os controles, embora em menor proporção (Tabela 1). Assim, por exemplo, ao estímulo “*Everyday he walks eight miles. Yesterday he _____*”, um dos indivíduos testados respondeu: “*Had a rest*”.

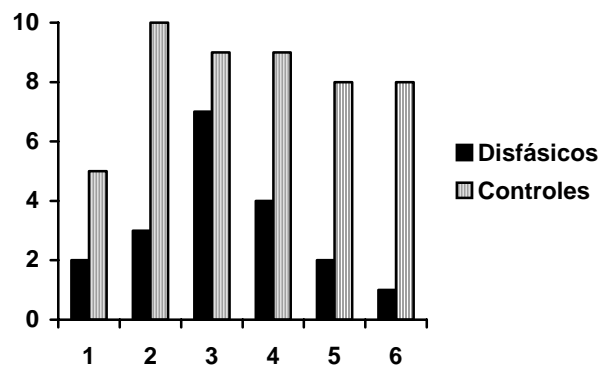


Figura 14. Resultados do teste de flexão verbal de 10 frases, mostrando diferença significativa entre os indivíduos afetados e os controles da família KE. Adaptado de Gopnik & Crago (1991).

	Tipo de erro											
	<i>Verbos não marcados</i>					<i>Verbos semanticamente relacionados</i>						
Disfásicos	2	1	3	0	4	5	4	5	1	2	0	1
Controles	0	0	0	0	0	0	2	1	0	0	0	1

Tabela 1. Frequência dos tipos de erros mais cometidos pelos indivíduos nos dois grupos investigados, pareados por idade: erros por omissão da marca de tempo (n=15, todos no grupo disfásico) e erros por substituição do verbo elicitado por um verbo semanticamente relacionado (n=13 no grupo disfásico versus n=4 no grupo controle). Adaptado de Gopnik & Crago (1991).

Essas duas estratégias – uso de formas não marcadas ou itens semanticamente, mas não morfologicamente relacionados – foram empregadas tanto nos testes de morfologia derivacional quanto nos de flexão verbal, gerando resultados semelhantes do ponto de vista estatístico. Esse achado empírico merece atenção, uma vez que constitui um indicativo de que, em termos de representação mental, podem ser mínimas as diferenças entre morfologia

derivacional e flexional. Na literatura, não há consenso quanto à atribuição desses processos morfológicos ao sistema computacional ou ao léxico (para discussão ver Alegre & Gordon, 1999a, 1999b). Uma observação pessoal das autoras parece importante quanto à postura dos indivíduos frente a esses dois testes:

“In both of these tests the dysphasics seemed not to be able to understand that the point of the test was to manipulate an underlying grammatical rule, though the normals understood with no hesitation. Even when they were prompted with examples and with specific instructions they seemed unable to perceive that there was an underlying pattern that was governed by a general rule. In this context, their strategy of responding semantically was intelligent” (Gopnik & Crago, 1991:44).

Por fim, Gopnik & Crago (1991) demonstram que embora os indivíduos disfásicos não apresentem dificuldades na compreensão de pequenas narrativas, a produção dessas, a partir de seqüências de imagens (Fig. 15), impõe um desafio ao grupo. Um dos parâmetros avaliados a partir da história ilustrada na Figura 15 foi a habilidade de recorrer a anáforas (“*he*”, p.ex.) em oposição ao uso repetido do sintagma nominal completo (“*the boy*”, p.ex.) (Quadro 4). Os resultados mostraram um uso significativamente maior de sintagmas nominais completos na narrativa dos indivíduos disfásicos ($p=0,00001$).

Uma vez que os pronomes são elementos com marcação de traço na gramática, segundo a interpretação das autoras, esse achado também pode ser explicado pela “*feature-deficit theory*”. A idéia é que esses indivíduos lexicalizariam os elementos pronominais, de modo que “*he*” indicaria algo como ‘indivíduo do sexo masculino não especificado’. Ou seja, na gramática desses indivíduos, os pronomes são representados como itens lexicais, não-marcados, com um significado específico atribuído. Assim, testes em que os pronomes eram empregados de um modo direto (p.ex., associação de uma figura a uma frase contendo pronome, nas tarefas de compreensão gramatical) não representavam grande dificuldade, uma vez que podiam ser resolvidos com base tão somente no significado associado ao pronome.

Por outro lado, o uso da anáfora verdadeira exige uma compreensão da associação entre os traços no pronome e os traços no sintagma nominal ao qual ele se refere.



Figura 15. Estímulo para produção de narrativa curta de modo a avaliar a capacidade de representação verbal dos eventos ilustrados e, particularmente, o uso de anáforas na narração dos eventos retratados. Extraído de Gopnik & Crago (1991).

Teste de produção de narrativa curta – uso de pronomes, sintagmas e anáforas

<i>Disfásicos</i>	<i>Controles</i>
<p><i>The lady's pointing at the bird and the man is watching <u>her</u>. The man is climbing the tree looking at the birds in their nest. The lady is crying because the man what fall off the tree (pause) and the bird frowd away. The neighbors phone the ambulance because the man fall off the tree. The ambulance come along and put the man into the ambulance to the hospital. And <u>he's</u> got a broken leg in the hospital.</i></p>	<p><i>The woman's looking at the birds and <u>she's</u> saying to <u>him</u>. "Oh, they're lovely birds, but <u>they</u> got stuck in the tree." So <u>he</u> says, "Well. I'll get <u>them</u> down then." So <u>he</u> climbs up the tree and the branch breaks. So <u>she's</u> nearly in tears cause <u>she</u> swallowed the wrapping string, and the birds fly off and <u>they</u> fall to the ground cause <u>they</u> can't fly. So <u>he</u> falls on the ground and <u>he</u> breaks <u>his</u> leg. So <u>she</u> runs indoors to call an ambulance with all the neighbors outside the door. The ambulance comes. <u>They</u> get in the van and he's in the hospital, and the birds crying by the window.</i></p>

Quadro 4. Comparação entre o uso de sintagmas nominais completos (em itálico), pronomes e anáforas (sublinhados) em duas narrativas produzidas a partir do estímulo apresentado na Figura 15.

Uma última colocação interessante quanto ao estudo desenvolvido por essas autoras diz respeito à percepção subjetiva das mesmas sobre a atitude dos indivíduos testados durante as tarefas.

“There are long hesitations and qualifiers such as “I’m not very good at this.” They respond as if they are solving a very difficult problem: as if they were asked to make a grammatical judgement in a language they were only somewhat familiar with. The normals, on the other hand, have absolutely no hesitation in responding to these items” (Gopnik & Crago, 1991:46).

Para tornar objetiva essa avaliação, as autoras estimaram o tempo transcorrido entre a pergunta e a resposta em três tarefas: morfologia derivacional, flexão verbal e estrutura argumental. Não surpreendentemente, nas duas primeiras, o tempo de resposta dos disfásicos foi praticamente o dobro dos controles, enquanto na terceira não houve diferença significativa entre os grupos.

Em suma, as autoras constroem um argumento em favor de um déficit seletivo gramatical, evidenciado por dificuldades específicas em tarefas de morfologia derivacional e flexional. Especulativamente, sugerem que a dificuldade seletiva na manipulação dos traços na língua teria origem num processo alterado de aquisição da linguagem, uma vez que a criança disfásica, diferentemente da criança normal, seria incapaz de construir um paradigma mental para novos casos a partir de dados anteriormente observados. Dito de outro modo, por serem incapazes de perceber as regularidades subjacentes à língua, não construiriam um modelo, nem seriam capazes de compreender que cada instância de uso da língua é, em si mesma, uma representação de determinadas propriedades gerais recorrentes.

Do que foi discutido até aqui, fica evidente que Gopnik & Crago (1991) propuseram uma mudança radical na descrição do déficit de linguagem da família KE em relação às observações de Hurst et al. (1990), mudança essa evidenciada pela própria reclassificação do distúrbio, que passou de apráxico para disfásico. Os dados gerados pelas autoras e publicados

em detalhe na *Cognition* de 1991 já haviam sido antecipados numa comunicação breve divulgada na revista *Nature*, no ano anterior, sob o também sugestivo título “*Feature-blind grammar and dysphasia*” (Gopnik, 1990). Nessa comunicação, ainda com resultados preliminares dos testes discutidos no artigo de 1991, Gopnik destaca que o déficit observado na família londrina estaria centrado em torno das regras gramaticais para paradigmas morfológicos flexionais regulares (especificamente, marcação de passado “-ed” e formação de plural “-s”), mantendo-se preservadas as habilidades lexicais de memorização de palavras e significados, o que justificaria o melhor desempenho dos KE em testes de flexão de irregulares (“go → went”, “mouse → mice”).

A forma como Gopnik analisa os déficits da família KE fundamenta-se, claramente, na concepção chomskyana de organização das línguas em torno de um sistema cognitivo bipartido entre léxico e sistema computacional, os quais teriam relativa autonomia do ponto de vista da sua realização no cérebro, de modo que poderiam ser seletivamente lesados. Os dados gerados por Gopnik & Crago vieram ao encontro da expectativa de diversos pesquisadores, tributários de Chomsky, que viram na família KE a possibilidade de dar maior concretude aos seus postulados teóricos. Em 1991, Pinker publica na revista *Science* um artigo intitulado *Rules of Language*, no qual discute a teoria da modularidade da mente de Fodor (1990) e defende, em particular, a existência de um domínio gramatical específico, dissociado de outras habilidades cognitivas.

Para tanto, Pinker vale-se, além dos resultados empíricos obtidos por Gopnik sobre a família KE, que sugerem um déficit restrito à gramática e à aplicação de regras, dos dados obtidos por outros pesquisadores sobre indivíduos com uma síndrome genética peculiar, que se tornaria um grande foco de atenção para pesquisadores das relações entre cérebro e linguagem – a síndrome de Williams (SW). Ao esboçar, ainda que de modo não-explicito, um contraste entre os dados de linguagem da família KE – tomada como paradigma de SLI – e os

indivíduos com a SW, Pinker lança a pedra fundamental de uma acirrada discussão em torno da dupla dissociação da linguagem.

A SW é uma condição genética relativamente rara, resultante de uma microdeleção no braço longo do cromossomo 7 (7q11.23) envolvendo uma série de genes situados nessa região. Mais de 15 genes já foram associados à síndrome, sendo duas as deleções mais comumente relatadas: a do gene *ELN1*, que codifica a proteína elastina e seria responsável por defeitos vasculares, e a do *LIMK1*, gene que se expressa no sistema nervoso central e estaria mais diretamente relacionado ao perfil cognitivo da SW.

Já durante a infância, as crianças com SW apresentam uma série de manifestações clínicas, que incluem aumento dos níveis de cálcio no sangue (hipercalcemia), retardo no crescimento e no desenvolvimento motor, malformações cardíacas, anomalias renais, traços faciais característicos e um nível variado de déficit cognitivo, com QI médio de 60, compatível com retardo mental leve (Morris & Mervis, 2000). No entanto, o que chamou a atenção dos pesquisadores para a SW foi o seu perfil cognitivo único, que mistura pontos fortes e fracos em aspectos específicos da cognição: o que se observa é uma proficiência lingüística notadamente superior ao desenvolvimento de outras habilidades cognitivas, em especial, ao processamento de informações visuo-espaciais (Fig. 16). Além disso, sabe-se que, para esse grupo de indivíduos, quanto melhor o desempenho em testes de proficiência lingüística, maior a discrepância entre as habilidades verbais e as não-verbais (Jarrold, 1998).

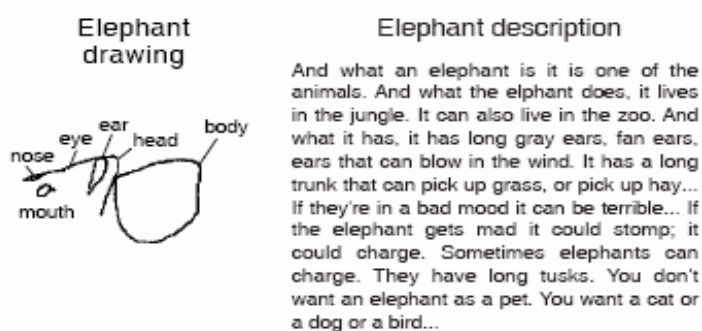


Figura 16. Comparação entre o desenho e a descrição oral de um elefante por um adolescente com síndrome de Williams. Extraído de Bellugi et al. (1999).

Esse perfil heterogêneo de habilidades tornou-se bastante evidente à medida que crianças com SW foram sendo avaliadas não apenas em comparação a crianças com desenvolvimento normal, mas em comparação a crianças com retardo mental de outras etiologias. Particularmente, o contraste com crianças com a síndrome de Down terminou por revelar uma curiosa imagem em espelho entre as duas síndromes no que tange ao processamento lingüístico e visuo-espacial (Fig. 17). Enquanto as crianças com síndrome de Williams apresentam uma fala fluente e gramatical, amparada por um rico vocabulário, mas demonstram imensa dificuldade com o processamento e a representação visuo-espacial, as crianças com síndrome de Down comportam-se de modo inverso, tendo habilidades visuo-espaciais francamente superiores às suas habilidades lingüísticas (Vicari, 2004).

Mais do que uma dissociação entre as habilidades cognitivas, a SW também evidencia uma dissociação dentro do próprio domínio lingüístico. O desenvolvimento do vocabulário seria um ponto forte dos indivíduos com SW, assim como a fluência verbal, enquanto as habilidades morfossintáticas seriam um ponto fraco na síndrome. Revelando um importante contraste com os membros da família KE – que parecem “desconhecer” as regras de formação de passado – um dos primeiros achados indicou que, dentre os poucos erros cometidos pelos indivíduos com SW, observa-se uma tendência à supergeneralização da regra de formação do passado (Bellugi et al., 1990), fenômeno também observado, durante um determinado período, no desenvolvimento normal da linguagem.

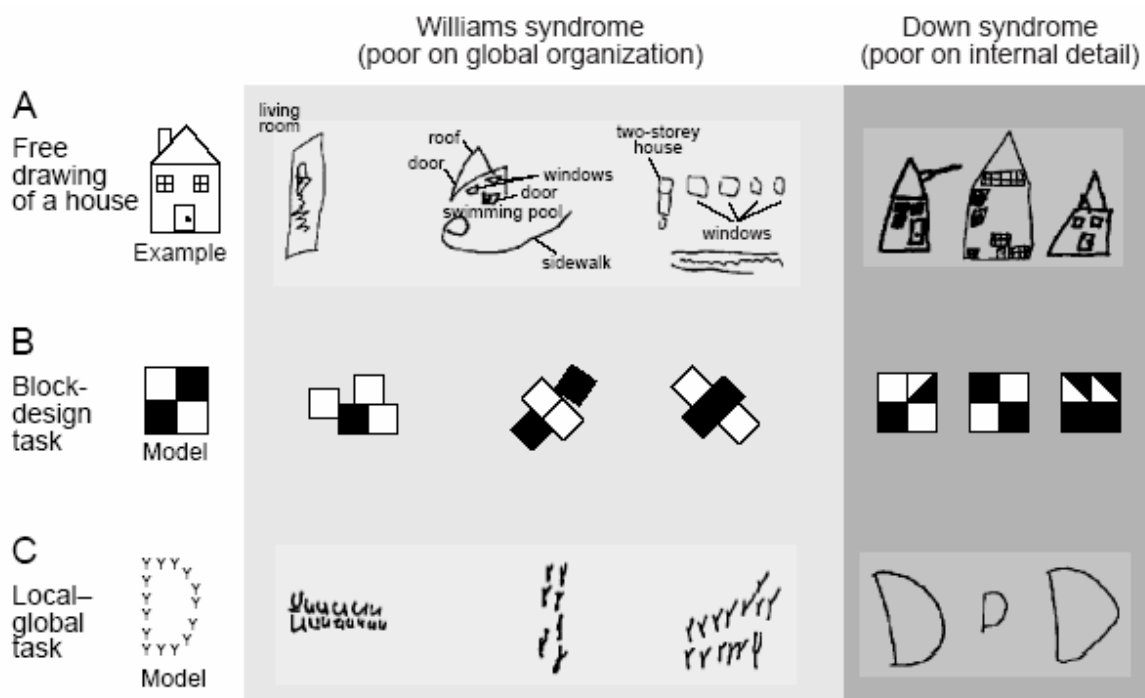


Figura 17. Comparação entre as habilidades visuo-espaciais de crianças com síndrome de Williams e síndrome de Down (extraído de Bellugi et al., 1999).

Estudos posteriores à publicação de Pinker confirmaram os achados preliminares de Bellugi, indicando que indivíduos com SW atingem altos níveis de acerto quando solicitados a flexionar verbos regulares, mas não alcançam o mesmo desempenho com verbos irregulares (Clahsen & Almazan, 1998, 2001). Comparado a crianças com desenvolvimento normal, o grupo SW regulariza a flexão de verbos irregulares de forma significativamente mais freqüente, optando pelo acréscimo de *-ed* em contextos nos quais o grupo controle raramente opta pela regularização. Esses achados foram replicados por Krause & Penke (2002) para a língua alemã e por Pleh et al. (2003) para o húngaro, demonstrando ainda que esse fenômeno não se restringe à flexão verbal, mas também à formação do plural.

Desse modo, os achados relativos à formação do passado e do plural na SW – por si só – oferecem uma evidência valiosa em favor da hipótese desenvolvida por Pinker (1991, 2002), baseada no modelo teórico de Chomsky, de que a linguagem está organizada sob a forma de um sistema dual, isto é, de um sistema computacional (responsável pela aplicação de regras como “acrescentar *-ed*”) e outro lexical (encarregado, entre outras coisas, da

recuperação de palavras do dicionário mental, como na flexão plural *mouse* → *mice* e na recuperação do passado irregular *eat* → *ate*). Mais do que isso, as habilidades contrastantes em relação ao paradigma flexional do inglês na SW e na família KE (SLI) sustentam uma possível independência entre esses dois subsistemas, sendo este, especulativamente, um modelo único de dupla dissociação entre regras gramaticais e processos lexicais (Quadro 5).

	Síndrome de Williams	Família KE (SLI)
Regras gramaticais	preservado	comprometido
Processos lexicais	comprometido	preservado

Quadro 5. Esquema ilustrativo da dupla dissociação entre regras gramaticais e processos lexicais na Síndrome de Williams e no déficit específico de linguagem (SLI), considerado a partir dos dados da família KE. Na dupla dissociação, que sugere uma independência entre módulos funcionais do sistema nervoso, uma determinada lesão prejudica o desempenho de uma função “A” deixando o desempenho de outra função “B” intacto, ao passo que um segundo tipo de lesão produz o efeito oposto, i.e., o desempenho da função “A” está intacto enquanto o da função “B” está prejudicado.

As tentativas de apropriação e conformação dos dados da SW e da família KE para a sustentação de hipóteses em favor de um domínio específico da linguagem geraram, naturalmente, uma resposta vigorosa por parte dos pesquisadores contrários a esse modelo de organização da linguagem. Pouco depois da primeira publicação de Gopnik, Vargha-Khadem & Passingham (1990) – pesquisadores alocados no *The Great Ormond Street Hospital for Children*, de Londres, onde a família KE vinha sendo estudada desde 1988 – publicam, também na *Nature*, uma resposta à análise de Gopnik: “*Gopnik has focused on only one aspect of the disorder, and it is inaccurate to conclude that the speech and language problem stems simply from an impairment in manipulating feature-markers*” (p.226).

A afirmação de Vargha-Khadem foi embasada nas conclusões de um estudo liderado por ela, cujos resultados foram divulgados em 1995. Recuperando a essência da caracterização fenotípica inicial de Hurst et al. (1990), Vargha-Khadem et al. (1995) assinalam que “*the inherited disorder does not affect morphosyntax exclusively, or even primarily; rather, it affects intellectual, linguistic, and orofacial praxic functions generally*”

(p.930). Nos diversos testes realizados para avaliação de aptidões fonológicas, lexicais e gramaticais – quando membros afetados e não-afetados foram pareados por idade – revelou-se um desempenho uniformemente inferior entre os afetados e, particularmente, não foram confirmados os achados de Gopnik & Crago quanto à suposta dissociação entre morfologia regular e irregular.

Vargha-Khadem et al. (1995) testaram 21 membros da família KE (13 afetados, 8 não-afetados), os quais, à exceção de um, eram os mesmos analisados por Gopnik & Crago (1991)³. A média de idade no grupo afetado foi de 24,4 anos (variando de 6 a 75) e, no grupo controle, de 14,1 (variando de 8 a 21). Nessa conformação, o grupo controle obteve resultados melhores em 11 dos 13 testes, com a diferença entre os dois grupos apontando significância estatística (em apenas dois não houve diferença significativa: nomear objetos a partir de figuras e apontar figuras a partir de palavras faladas pelo entrevistador). Quando os resultados da segunda geração foram excluídos da análise, segundo os autores, para “melhor equiparar” os indivíduos testados quanto à idade, apenas um dos 13 testes (nomear objetos) não mostrou diferença significativa entre os grupos. Esse ligeiro aumento na discrepância entre os dois grupos, com a eliminação da segunda geração, poderia ter sido esperado, tendo em vista os relatos consistentes de que o déficit de linguagem torna-se menos marcante a partir da puberdade. Se o objetivo dos autores era um melhor pareamento por idade, por que os dados da avó, a única que não recebeu acompanhamento fonoaudiológico, não foram excluídos? Atente-se ao fato de que, à exceção desta, todos os demais sujeitos testados neste estudo eram da primeira geração, portanto, aqueles com déficit oromotor mais evidente.

Particularmente relevante na análise de Vargha-Khadem et al. (1995) foram os resultados do teste de produção de formas verbais flexionadas com marca de tempo. A fim de

³ Embora Gopnik & Crago (1991) tenham avaliado 22 membros da família KE, entre afetados e controles, apenas os dados de 12 foram incluídos na análise dos resultados. Logo, a amostra é diferente daquela de Vargha-Khadem et al. (1995).

avaliar a capacidade dos indivíduos de alterar a forma de verbos do presente para o passado e vice-versa, foram apresentadas frases como “*Everyday I wash my clothes. Yesterday I _____ my clothes*”. Do total de 40 frases, os resultados apontaram uma média de acertos de 37,43 entre os não-afetados contra 19,91 entre os afetados, reafirmando a dificuldade com a morfologia flexional, mas sem evidenciar qualquer diferença no desempenho com verbos regulares ou irregulares. Além disso, novamente contrariando a análise de Gopnik e Pinker, os pesquisadores indicaram que 41% dos erros cometidos pelos indivíduos afetados foram, justamente, de supergeneralização da regra “*V + -ed*”.

A justificativa dos autores para essa discrepância de resultados em relação àqueles de Gopnik & Crago (1991) é pouco esclarecedora: a bateria de testes dessas autoras teria empregado um número menor de verbos e frases. Uma análise mais detalhada do conjunto de resultados de Vargha-Khadem et al. (1995) parece mais reveladora. Na comparação dos resultados dos dois grupos, as avaliações que apresentaram teste de significância bastante expressivo tinham comandos como “*Repeat this word*”, “*Read this nonword*” e até mesmo “*Repeat this nonword exactly as I say it*” (grifo meu). Por outro lado, os testes que evidenciaram diferença menos marcante entre os grupos tinham comandos como “*Is this a real English word?*”, “*Show me the picture for this word*” e “*Tell me the name of the object in this picture*”. Claramente, o escore atribuído aos testes do primeiro grupo de comandos perpassa critérios fonológicos – ponto de déficit evidente e reconhecido entre os membros afetados da família KE – enquanto as respostas para o segundo grupo de comandos incluem sim/não, apontar e nomear (sendo o conteúdo semântico da resposta, neste último caso, mais importante que o critério fonológico). Assim, os escores do primeiro grupo são muito possivelmente entrevistador-dependente e, portanto, mais subjetivos, com a qualidade fonológica da resposta exercendo um peso substancial sobre o julgamento da adequação ou não da resposta.

Outro achado importante desse estudo assinala uma dificuldade significativa dos membros afetados da família KE com o manejo de pseudopalavras – as chamadas *nonwords*, seja em sua leitura, escrita, repetição, deleção/adição de fonemas ou pronúncia, dificuldade esta que viria a ser enfatizada por estudos posteriores. Testes com pseudopalavras são freqüentemente utilizados para avaliar a alça fonológica da memória de trabalho, sistema dedicado ao armazenamento e à manipulação de informações durante um curto período de tempo (a memória de trabalho é empregada, por exemplo, para reter um número de telefone até que ele seja discado). A alça fonológica seria composta por dois componentes: um arquivo fonológico, capaz de guardar informações provenientes da fala, que desaparecem ao fim de alguns segundos, e um processo articulatorio, capaz de restaurar as informações perdidas, levando-as novamente para o arquivo fonológico através de um processo de ensaio subvocal (Baddeley, 2003). Diversos estudos sugerem que esse subsistema da memória de trabalho seria importante para o desenvolvimento de habilidades de leitura na infância, para a compreensão da linguagem falada e para a aquisição de vocabulário (para revisão ver Silva Nunes, 2001), havendo uma correlação entre a capacidade de repetir pseudopalavras e o tamanho do vocabulário (Munson et al., 2005). Desse modo, esses achados podem ser interpretados como indícios de um déficit de memória de trabalho (atestado também por um desempenho inferior aos controles no teste de rememoração de lista de números – *digit span*), com repercussões sobre o processo de aquisição da linguagem.

Tanto os testes de morfologia flexional quanto de derivacional revelaram resultados semelhantes aos de Gopnik & Crago (1991), confirmando que a produção e a compreensão de marcadores morfológicos está prejudicada nos indivíduos afetados, tanto no que diz respeito às palavras existentes no léxico da língua inglesa quanto às pseudopalavras (Tabela 2).

Teste	Escore máximo	Escore afetados	Escore controles	p
Produção de marcadores morfológicos*				
Palavras	20	14,17 ± 2,86	19,29 ± 0,76	<0,001
Pseudopalavras	20	6,83 ± 2,76	16,00 ± 1,63	<0,001
Julgamento de marcadores morfológicos‡				
Palavras	24	14,20 ± 5,15	20,86 ± 2,19	0,440
Pseudopalavras	24	9,22 ± 3,67	14,14 ± 5,01	0,029

Tabela 2. Seleção de resultados, apresentados como média ± desvio-padrão, comparando o desempenho em testes com palavras e pseudopalavras. * Palavras: “*This creature is smaller than this one, but this creature must be the [smallest].*” Pseudopalavras: “*This creature is ponner than this one, but this one must be the [ponnest].*” ‡ Qual das seguintes frases está correta? Palavras: “*Planes are faster than trains, or planes are fastest than trains?*” Pseudopalavras: “*Planes are donker than trains, or planes are donkest than trains?*”. Dados extraídos de Vargha-Khadem et al. (1995)

A aplicação de testes de QI ampliou a caracterização fenotípica dos KE. Tanto na análise do QI verbal quanto do não-verbal observou-se uma diferença estatisticamente significativa entre os grupos, conforme ilustrado na Tabela 3. Para referência, observe-se que um QI de aproximadamente 50-55 até 70 é compatível com retardo mental leve, classificação que congrega 85% dos indivíduos com o transtorno (APA, 1995). Como comentado anteriormente, o retardo mental é um dos critérios de exclusão para o diagnóstico de SLI. Seis indivíduos afetados (e um não-afetado) apresentaram um escore inferior a 85 na avaliação de QI não-verbal, considerado na literatura como ponto de corte para exclusão do diagnóstico de SLI, já que indicaria um déficit cognitivo mais amplo, não restrito à linguagem (Tomblin et al., 1996). É nessa linha que os autores interpretam esse achado: “*the cognitive impairment of the affected family members is not confined to morphosyntax. Rather, it appears to extend to the verbal domain in general, and, indeed, is just as great in the nonverbal domain*” (p. 932).

	Afetados	Controles	P
QI verbal	75 (59-91)	94 (82-111)	<0,001
QI não-verbal	86 (71-111)	104 (84-119)	0,004

Tabela 3. Comparação dos resultados nos testes de QI entre os grupos caso e controle. Foram utilizadas na avaliação as escalas WISC, *Wechsler Intelligence Scale for Children – Revised*, (5 a 16 anos) e WAIS, *Wechsler Adult Intelligence Scale – Revised*, conjunto de instrumentos disponível em diversos idiomas e empregado em larga escala para avaliação global da inteligência. Elaborada a partir dos dados de Vargha-Khadem et al. (1995).

Contudo, é interessante observar aqui uma possível sobreposição parcial entre déficit lingüístico e baixo QI não-verbal, retomando também a observação de que distúrbios de linguagem co-ocorrem com retardo mental em diversas síndromes genéticas. Com a descoberta da mutação no *FOXP2*, a relação entre a família KE e as crianças com SLI pode ser investigada também sob uma perspectiva genotípica. Embora a mutação específica G→A no éxon 14, observada na família KE, não tenha sido encontrada em outros indivíduos diagnosticados com SLI (Newbury et al., 2002; SLI Consortium 2002; O'Brien et al., 2003), marcadores de regiões adjacentes ao *FOXP2* mostraram uma forte associação com esse tipo de déficit, sugerindo que mutações fora da região codificadora do gene, envolvendo regiões regulatórias, possam estar envolvidas na etiologia de um fenótipo de SLI mais leve do que aquele observado na família KE (O'Brien et al., 2003).

Os critérios de inclusão empregados na seleção de crianças com SLI para a realização desses estudos consistiam, basicamente, em testes de linguagem e QI. Newbury et al. (2002) e o SLI Consortium (2002) excluíram de suas amostras crianças com QI não-verbal <80 e não encontraram nenhuma associação. Quando esse critério de exclusão foi rebaixado (ou relaxado) para um QI não-verbal <70, O'Brien et al. (2003) encontraram associação com marcadores adjacentes ao *FOXP2*. Embora essa relação possa ser verdadeira, é possível, todavia, que a inclusão de crianças com um QI não-verbal baixo tenha atuado como fator confundidor da análise genética.

Retomando o estudo de Vargha-Khadem et al. (1995), foi demonstrado, por fim, que a família KE apresenta uma dificuldade substancial na execução de movimentos orofaciais não-articulatórios, ou seja, não-lingüísticos, estendendo a classificação da dispraxia para oral, e não apenas verbal. Como mencionado anteriormente, foram realizados tanto testes que exigiam o cumprimento de comandos verbais quanto testes de imitação, a fim de garantir que a apresentação verbal da tarefa, por possíveis dificuldades de compreensão, não estivesse

interferindo nos resultados. Foram analisados movimentos simples (p.ex., abrir a boca), complexos (p.ex., protruir o lábio e a mandíbula), simultâneos (p.ex., abrir a boca, pôr a língua para fora e vocalizar) e seqüenciais (p.ex., cerrar os lábios, depois abrir a boca e, então, pôr a língua para fora) (Fig. 18).

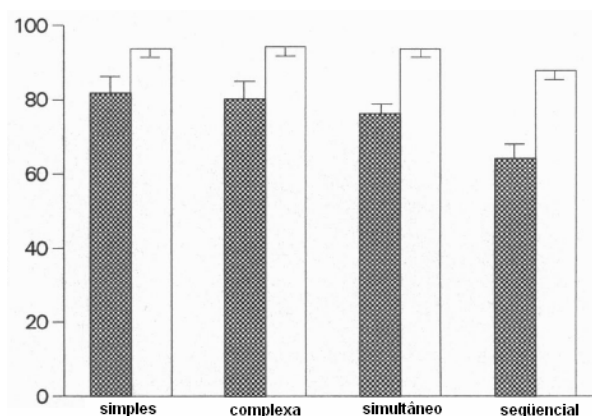


Figura 18. Resultados de Vargha-Khadem et al. (1998), expressos em percentual de acertos, para testes de praxia oral da família KE (barras escuras), revelando dificuldades substanciais na execução de movimentos orofaciais não-articulatórios (i.e., não-lingüísticos), em especial, movimentos seqüenciais complexos.

Em 1998, quando a região SPCH1 já havia sido identificada (Fisher et al., 1998), Vargha-Khadem et al. apresentam um novo trabalho, no qual outros testes reafirmam esse déficit de articulação seqüencial e praxia orofacial, dando-lhe primazia na descrição do fenótipo da família KE, ao considerá-lo como “o déficit nuclear” dos membros afetados. Os autores sugerem ainda que a dispraxia verbal dos KE – que tornaria a fala desses indivíduos praticamente incompreensível ao ouvido não-treinado – seria, de fato, uma consequência da dispraxia orofacial, manifesta através de uma relativa imobilidade da metade inferior da face (em especial, do lábio superior), afetando também os movimentos faciais não relacionados à fala. Ainda no que diz respeito à ligação entre os déficits, os autores observam que embora os escores individuais nos testes de repetição de palavras e de pseudopalavras estejam correlacionados, não há correlação entre estes e os escores nos testes de praxia orofacial. Nesse sentido, comentam:

The results indicate that even though both the orofacial movement and speech repetition tests are probably tapping into the same

core deficit—i.e., both are measures of oromotor coordination and every affected individual is impaired on both—there are also important differences between the two, with speech repetition requiring a far greater oromotor range, complexity, and precision than the other (Vargha-Khadem et al. 1998:12697).

Esses resultados parecem sugerir, assim, que o déficit central da família KE relaciona-se a um componente de motricidade oral mais amplo, não relacionado exclusivamente à fala. Além dos resultados dos testes de praxia orofacial, que fornecem uma medida quantitativa do déficit em sua face não-verbal, o próprio teste de repetição de palavras – ao mostrar uma diferença muito pequena entre o desempenho com palavras existentes e pseudopalavras – sugere que o déficit da família KE não esteja relacionado diretamente ao subsistema fonológico da memória de trabalho; se assim fosse, a repetição de pseudopalavras deveria impor dificuldades adicionais.

A idéia de estabelecer um déficit nuclear ou central para o fenótipo dos membros afetados da família KE passa obrigatoriamente pela determinação de uma característica que preencha, no mínimo, dois critérios: (i) o déficit em questão é observado em todos os membros afetados e (ii) não é observado entre qualquer um dos membros não-afetados. Nesse sentido, Vargha-Khadem et al. (1998) trazem uma contribuição importante ao afirmar que, embora os escores médios do grupo afetado sejam consistentemente inferiores aos dos não-afetados na ampla maioria dos testes de fala, linguagem e motricidade oral, individualmente apenas três testes mostram diferenças claras entre afetados e não-afetados: repetição de palavras, repetição de pseudopalavras e movimentos orofaciais (seqüenciais e simultâneos). Nesses, não se observa qualquer sobreposição entre os resultados individuais dos membros afetados e os do grupo controle, de modo que os resultados desses testes poderiam ser usados como uma medida objetiva para distinguir se determinado membro da família tem ou não o problema.

Em 1999, Ullman & Gopnik divulgam os resultados de uma abrangente bateria de testes investigando especificamente as habilidades morfológicas flexionais de 10 membros da família KE, dos quais sete eram afetados, com idades variando entre 10 e 77 anos (da primeira à terceira geração). Foram utilizados 56 verbos, divididos entre regulares, irregulares e pseudoverbos com padrão regular e irregular. Os autores observaram um desempenho significativamente inferior dos membros afetados (comparados a controles não-aparentados, pareados por idade) nas seguintes tarefas: formação de passado regular de pseudoverbos (p.ex., *plam* → *plammed*) ($p < 0,001$); formação de passado regular ($p=0,004$) e irregular ($p=0,005$) de palavras existentes. A análise do resultado dos testes volta a reafirmar um padrão de não-aplicação (ou de aplicação irregular) da regra de sufixação (-ed) na tarefa de produção de passado regular, sugerindo uma incapacidade de supergeneralização da regra para novos casos (“*failure to overregularize*”).

Alcock et al. (2000) confirmaram que o desempenho dos membros afetados da família KE era inferior ao de indivíduos normais, pareados por idade, em testes de movimentos orais, fossem eles de natureza verbal ou não. O único teste no qual esses indivíduos não mostraram déficit significativo foi na execução de movimentos orais simples, que incluíam um único grupo de músculos (como, p.ex., “faça um som ‘ah’”, empregando apenas as cordas vocais). Já nos testes incluindo mais de um grupo muscular, assim como naqueles demandando uma execução de movimentos combinados (seqüenciais ou simultâneos), os resultados dos membros afetados da família KE foram significativamente inferiores. Uma vez que a fala exige um controle oromotor rápido e preciso, os autores sugerem que a origem do déficit articulatorio da família KE estaria enraizada nessa dificuldade generalizada na coordenação da musculatura oral para movimentos complexos.

Watkins et al. (2002a) conduziram uma bateria de testes com 13 membros afetados da família KE (média de idade 25,3 anos) em uma nova tentativa de definir o déficit nuclear

desses indivíduos. Desta vez, os resultados foram comparados com dois grupos controle: membros não-afetados da família (média de idade 17,1 anos) e indivíduos com afasia adquirida (média de idade 68,9 anos)⁴. Além de confirmar um QI verbal (média 74,1) e não-verbal (média 83,2) reduzido entre os membros afetados, o estudo demonstrou que o teste que melhor discrimina indivíduos afetados e não-afetados da família KE é um teste de repetição de pseudopalavras com um padrão complexo de articulação (encontros consonantais).

O teste de repetição de palavras, com articulação simples e complexa, mostrou que o grupo não-afetado tinha um desempenho significativamente superior ao grupo afásico e este, um desempenho significativamente superior ao dos afetados. Os afásicos mostraram um desempenho significativamente melhor na repetição de palavras em relação às pseudopalavras, mas este efeito, que sugere um prejuízo da memória de trabalho, mais uma vez, não foi observado entre os afetados, cujos resultados foram igualmente inferiores aos grupos controle em ambos os testes.

De modo geral, o grupo de afetados da família KE e o grupo de afásicos apresentaram um resultado semelhante, significativamente inferior aos não-afetados, em todos os testes verbais. As tarefas avaliadas por Watkins et al. (2002a) incluíram, além de repetição de palavras, nomeação de figuras, decisão lexical, recepção gramatical (associar uma de quatro figuras a frase, sintagma ou palavra correspondente) e fluência verbal (produzir oralmente, em dois minutos, o maior número de palavras começando com as letras “f” ou “m” e pertencendo às categorias “frutas” e “animais”). A forte semelhança entre o desempenho de afásicos e indivíduos afetados da família KE em tantas tarefas é curiosa, revelando que o estado atual das habilidades lingüísticas de ambos os grupos é semelhante, muito embora cada um tenha vivenciado um processo de aquisição diferente, com os afásicos apresentando um curso normal de desenvolvimento da linguagem previamente ao acidente vascular encefálico.

⁴ Os 11 indivíduos que compõe essa amostra têm um diagnóstico de afasia expressiva (de Broca), secundária a acidente vascular encefálico no hemisfério esquerdo.

A ênfase dada por Vargha-Khadem e seus colaboradores no estabelecimento dos aspectos oromotores como centrais ao déficit da família KE não deve, no entanto, fazer crer que o déficit morfossintático enfatizado por Gopnik e colaboradores não seja também significativo. Os achados de Watkins et al. (2002a) reforçaram a existência de uma dificuldade desses indivíduos com os aspectos morfológicos da língua, tanto flexionais quanto derivacionais, sendo, neste caso, a dificuldade mais evidente quando o teste vale-se de pseudopalavras (Fig. 19).

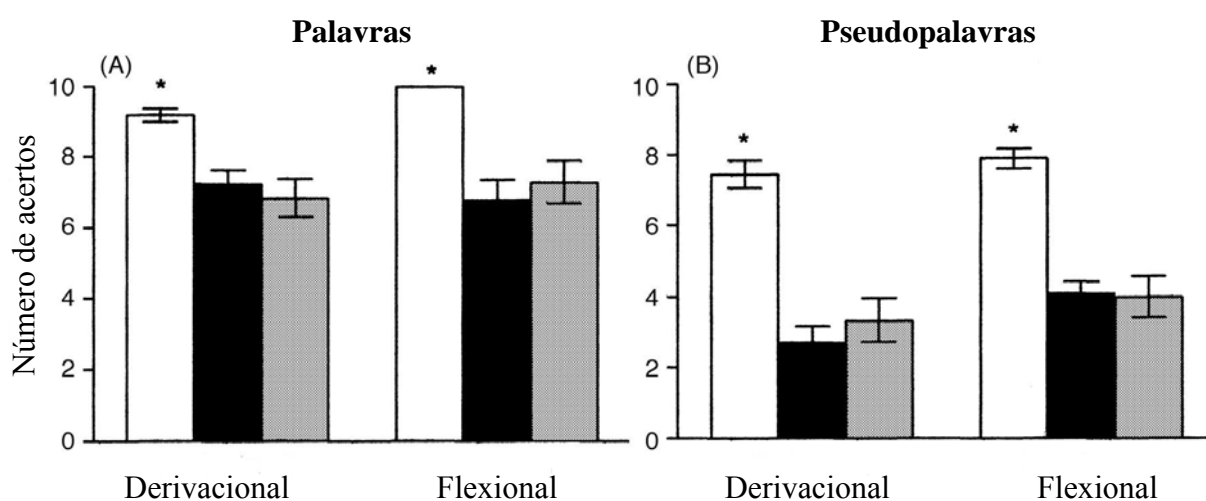


Figura 19. Comparação entre os resultados de membros não-afetados (branco), afetados (preto) e afásicos (cinza) em testes de morfologia derivacional e flexional utilizando palavras (A) e pseudopalavras (B). O asterisco indica diferença significativa entre o grupo não-afetado e os grupos afetado e controle, que não diferiram entre si. Adaptado de Watkins et al. (2002a).

O debate em torno da caracterização e da definição precisa do déficit nuclear dos membros afetados da família KE não se restringe, claramente, ao caso específico desta família, que representa, até hoje, o único caso relatado de um distúrbio no desenvolvimento da linguagem ocasionado por uma transmissão indubitavelmente monogênica (Fisher et al., 2003). É justamente a presença desse fator claramente genético na etiologia do déficit de linguagem dos KE que o torna um objeto de interesse aos defensores e, conseqüentemente, aos opositores do inatismo. A família KE representaria, assim, uma evidência material única em favor da hipótese chomskyana de que existiriam genes responsáveis pela organização de circuitos críticos subjacentes à GU – como comenta Pinker sobre essa família, ainda em 1994:

A síndrome mostra, portanto, que, no desenvolvimento do cérebro, deve haver algum padrão de eventos geneticamente determinados (neste caso, os eventos afetados por essa síndrome) especializado na urdidura das conexões da computação lingüística. E esses sítios de construção parecem envolver circuitos necessários para o processamento da gramática na mente, e não apenas para a articulação de sons da fala pela boca ou a percepção de sons da fala pelo ouvido (Pinker, 2002:414-5).

De modo bastante transparente – e, de certo modo, fomentando a discussão que se estabelece a partir daí – Pinker afirma em seguida:

Embora na infância os membros afetados da família sofressem de dificuldades na articulação da fala e apresentassem retardo no desenvolvimento da linguagem, a maioria deles superou os problemas de articulação, perdurando apenas os déficits relacionados com gramática (id., ibid.).

A leitura – consideravelmente díspare – do déficit central dos KE feita pelos grupos de Gopnik e Vargha-Khadem remonta às diferentes posições desses grupos em relação à natureza da linguagem e, mais especificamente, à hipótese do inatismo. De um modo geral, ambas as correntes concordam que o déficit afeta, ao mesmo tempo, a praxia oromotora e o conhecimento gramatical, mas a relação entre esses déficits permanece motivo de controvérsia. Na discussão em torno da primazia de um sobre o outro, contudo, a posição de Vargha-Khadem e colaboradores, sustentada também por Marcus & Fisher (2003), parece mais justificada: a dispraxia orofacial é o déficit central dos membros afetados da família KE porque é um aspecto proeminente do distúrbio e “nuclear” no sentido de que possibilita uma distinção clara e simples entre membros afetados e não-afetados. Apesar do esforço e do interesse de ambos os grupos em gerar dados na tentativa de estabelecer medidas objetivas das habilidades dos membros afetados, muitos dos testes aplicados esbarram em avaliações subjetivas sobre o desempenho dos indivíduos testados, como evidenciado pela discrepância de resultados na aplicação de testes muito semelhantes.

Apesar de toda a disputa em torno da caracterização do fenótipo, quando o genótipo da família KE foi revelado, em 2001, comprovando a tese de uma mutação monogênica, muito já se havia avançado no entendimento da relação entre mutação e efeito, especialmente na descrição dos efeitos lingüísticos e cognitivos da mutação do *FOXP2*. A descoberta do gene, contudo, abriu novas possibilidades investigativas, permitindo expandir a caracterização do fenótipo. A análise do padrão de expressão do gene no cérebro, correlacionada às peculiaridades da neuroanatomia e da neurofisiologia desses indivíduos, poderia oferecer uma evidência decisiva sobre a hipótese de que o *FOXP2* seria responsável, ao menos em parte, pela organização adequada da circuitaria cerebral necessária para o desenvolvimento normal da fala e da linguagem.

Por dentro do cérebro dos KE

A ressonância magnética funcional, os métodos neurofisiológicos, bem como a observação de pacientes neurológicos e indivíduos normais, constituem hoje a base do estudo da neurobiologia da linguagem. O cenário atual é muito diferente daquele no qual foram feitas as primeiras descobertas nesse campo, baseadas quase que exclusivamente na análise *post-mortem* do cérebro de pacientes que haviam perdido as habilidades de linguagem.

No século XIX, Pierre Paul Broca e Karl Wernicke entraram para a história ao relatar a descoberta de regiões específicas do cérebro relacionadas à linguagem. Em 1863, Broca descreveu, a partir do estudo de cérebros necropsiados, uma região na porção póstero-lateral do lobo frontal do hemisfério esquerdo que seria a responsável pela fala (área de Broca). Em 1874, Wernicke postulou que uma região posterior do lobo temporal do hemisfério esquerdo era responsável pela compreensão da linguagem (área de Wernicke); ligando essas duas áreas estaria um feixe de axônios denominado fascículo arqueado (Geschwind, 1970) (Fig. 20).

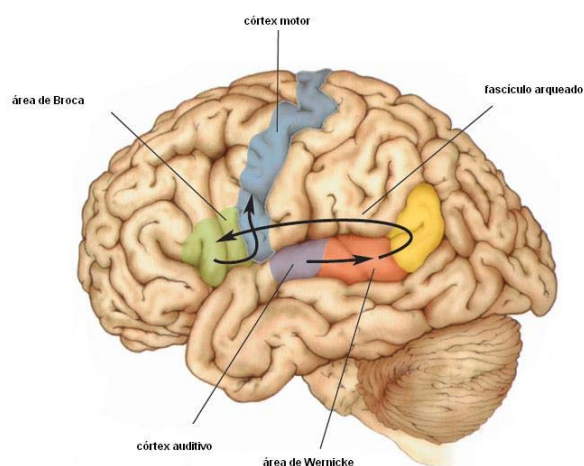


Figura 20. Localização das principais áreas do córtex cerebral associadas à linguagem. As setas na figura indicam a seqüência de regiões ativadas em uma tarefa de repetição de uma palavra falada: do pavilhão auditivo, os sons da fala são enviados na forma de sinais neurais ao córtex auditivo, não sendo compreendidos como palavras até seu processamento na área de Wernicke. Para repeti-las, sinais são enviadas à área de Broca, através do fascículo arqueado. Na área de Broca, esses sinais são convertidos em um código que instrui os movimentos necessários aos músculos da fala, que são ativados através do córtex motor. No caso da leitura, as imagens visuais são primeiramente transmitidas ao córtex visual (lobo occipital, porção mais posterior do encéfalo) e depois processadas na área de Wernicke para que a compreensão ocorra (Bear et al., 2002).

Hoje, contudo, sabe-se que a linguagem não envolve apenas essas áreas, mas sim uma complexa rede de circuitos neuronais funcionando como interconexões paralelas, e não apenas redes seriadas. As áreas de Broca e Wernicke permanecem como elementos-chave ao processo; todavia, mais recentemente, o papel de outras estruturas vem sendo investigado no processamento sensorial e motor da linguagem – não apenas no córtex, região associada aos processos cognitivos superiores – mas também em regiões subcorticais, as quais teriam uma função predominantemente integradora de sinais (Kandel et al., 1997).

No que tange ao caso específico da família KE, os dois elementos que se destacam como característicos do déficit – as habilidades gramaticais e a praxia oral, incluindo a articulação da fala – remetem, em princípio, a diferentes redes neurais. As habilidades gramaticais estão tipicamente ligadas a regiões do córtex frontal esquerdo, enquanto a praxia relaciona-se às funções motoras e está associada a regiões corticais e subcorticais.

A marcante dispraxia observada entre os membros afetados tem guiado a atenção dos pesquisadores à investigação de áreas relacionadas ao controle motor da fala. O planejamento

e o controle dos movimentos voluntários, como aqueles necessários à articulação da fala, envolvem diversas áreas do cérebro, dentre as quais destaca-se o córtex motor. A região destacada em azul na Fig. 20 não corresponde à totalidade do córtex motor (que se expande por áreas fronteiriças), representando exclusivamente o córtex motor primário (área M1). Essa região se caracteriza por apresentar uma somatotopia, ou seja, uma correlação entre regiões cerebrais e segmentos corporais. Assim, a estimulação elétrica de uma área específica do córtex motor primário desencadeia como resposta a contração de músculos em uma dada área do corpo, conforme o mapa apresentado na Fig. 21 (observe, p.ex., que o controle motor dos músculos da face e dos lábios está situado em uma porção lateral de M1).

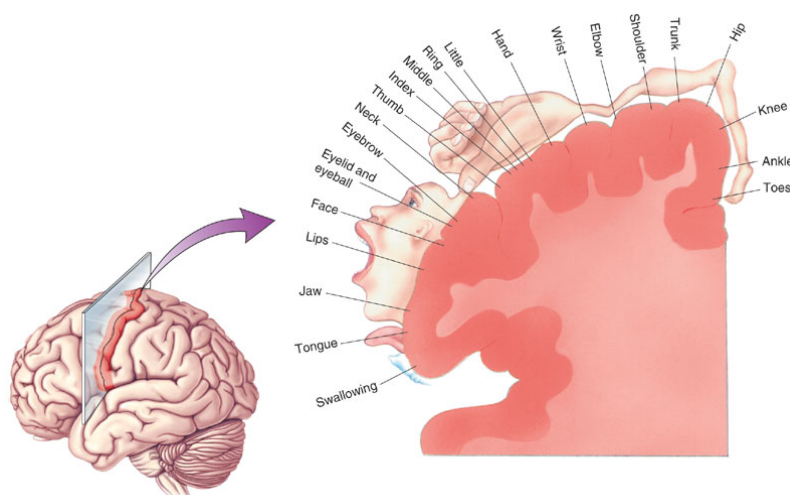


Figura 21. Representação esquemática da distribuição somatotópica das áreas motoras no córtex cerebral (M1). A descoberta dessa relação direta entre segmento cortical e segmento corporal foi feita pelo neurocirurgião Wilder Penfield, durante cirurgias de epilepsia, enquanto procurava mapear as áreas cuja função era tão importante que deviam ser poupadas do corte do bisturi. Extraído de Bear et al., 2002.

As áreas motoras do córtex cerebral estão, assim, relacionadas principalmente com o envio de sinais para a periferia (braços, pernas, boca, língua, etc.) para a execução do movimento. Essas regiões corticais, por sua vez, interagem com regiões subcorticais, dentre as quais destaca-se um conjunto de três estruturas conhecidas como núcleos da base: núcleo caudado, putâmen e globo pálido (Fig. 22). Embora participem de todas as etapas do movimento, os núcleos da base estão envolvidos principalmente nas etapas de programação do movimento, isto é, da seleção dos diferentes elementos que serão necessários à execução

do ato motor, tanto em termos temporais quanto espaciais. Acredita-se que os núcleos da base sejam responsáveis pelo controle de um acervo de reflexos inatos e automatismos adquiridos, que seriam seletivamente inibidos/desinibidos de acordo com o objetivo a ser atingido, “aliviando” o córtex cerebral da carga de atenção necessária à realização de movimentos específicos e permitindo a execução simultânea de diversos elementos do programa motor (Rodriguez, 1999).

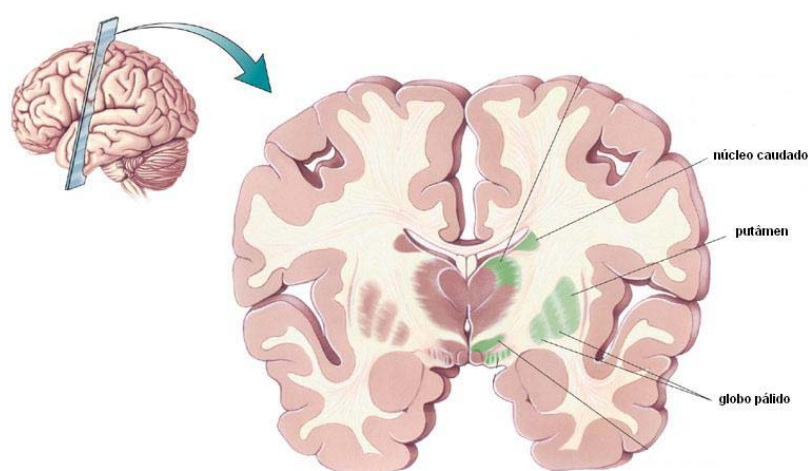


Figura 22. Corte coronal do cérebro humano. O córtex corresponde à região destacada no entorno. Em verde, os núcleos da base e regiões associadas. Feixes neuronais transitam entre as áreas corticais e subcorticais, mantendo-as em contato permanente. Extraído de Bear et al., 2002.

Estudos de neuroimagem permitem identificar o nível de atividade de uma região cerebral durante a execução de uma determinada tarefa. A partir desses estudos, é possível estabelecer quais são os padrões normais de atividade cerebral durante a leitura de um texto, a memorização de uma lista de palavras ou a realização de cálculos aritméticos, por exemplo. Um dos primeiros estudos de neuroimagem feitos com a família KE, ainda anteriormente à descoberta do *FOXP2*, mostrou que os membros afetados apresentavam alterações cerebrais não apenas funcionais, com padrões de ativação diferentes daqueles observados em indivíduos controles, mas também estruturais, com alteração no volume de corpos de neuronais em regiões específicas do cérebro (Vargha-Khadem et al., 1998).

Do ponto de vista funcional, os autores observaram que determinadas regiões do cérebro estavam subativadas ou superativadas durante a execução da tarefa descrita a seguir. Na condição de base, os indivíduos escutavam, através de fones de ouvido, palavras ditas em ordem reversa (i.e., lidas de trás para frente) e, como resposta, deviam dizer uma determinada palavra, previamente combinada. Na condição de teste, os indivíduos escutavam uma palavra normal e deveriam repeti-la em voz alta. Os testes foram feitos com dois afetados e quatro controles; para cada participante, foram geradas seis imagens da atividade cerebral em cada uma das duas situações, totalizando 72 imagens de PET *scan*⁵.

Os resultados revelaram que, durante a execução da tarefa, os indivíduos afetados mostravam menor nível de ativação de áreas do córtex associadas à atividade motora, incluindo as áreas relacionadas à face e à boca no hemisfério esquerdo (dominante para a linguagem), se comparados aos não-afetados. Por outro lado, as áreas de Broca, o núcleo caudado e o córtex pré-motor esquerdo – integrantes do sistema que, em sujeitos normais, é ativado para a produção de uma fala fluente – mostraram-se superativados nos afetados em relação aos controles.

Afirmar com certeza o que a superativação dessas áreas representa, em termos de processamento neural da linguagem, ainda não é possível. Basicamente, duas hipóteses podem ser consideradas: a primeira, de que se trata de um problema primário de superativação (a superativação como problema em si) que interfere na dificuldade de fala observada entre os membros afetados; alternativamente, a maior ativação pode ser pensada como um indício de um esforço extra, ou seja, de que a realização dessas tarefas impõe uma demanda adicional de

⁵ O PET *scan* ou Tomografia por Emissão de Pósitrons é uma tecnologia computadorizada que utiliza glicose radioativamente marcada para medir a atividade metabólica cerebral. Como a glicose é normalmente o “combustível” do cérebro, quanto mais ativa for uma dada parte do cérebro durante uma atividade, maior será a quantidade de glicose metabolizada. Um *scanner* com um conjunto de detectores de radiação mapeia o cérebro, localiza a radioatividade e manda esses dados para um computador que produz uma imagem bidimensional colorida, na qual as cores mais vivas indicam maior nível de atividade cerebral.

recursos para o processamento da linguagem nos indivíduos afetados da família KE (Vargha-Khadem et al., 1998). De qualquer modo, esses estudos permitem selecionar ou circunscrever “áreas candidatas” para a identificação da localização dos efeitos da mutação do *FOXP2* no cérebro humano.

Na busca por uma alteração estrutural que complementasse esse achado funcional, os autores demonstraram ainda que – embora bastante sutis – existem diferenças morfológicas entre o cérebro de afetados e não-afetados. A quantidade de substância cinzenta (representando, basicamente, o volume de corpos de neurônios) mostrou ser diferente em diversas regiões corticais e subcorticais dos 17 membros testados da família KE, quando comparados a controles pareados por idade.

O perfil fenotípico dos KE – com manifestação do distúrbio na infância e persistência, ainda que atenuada, na vida adulta – sugere que, do ponto de vista estrutural, a patologia desses indivíduos deva ser bilateral (i.e., observada tanto no hemisfério direito quanto no esquerdo)⁶. Além disso, de acordo com a hipótese de trabalho de Vargha-Khadem e colaboradores, a estrutura anatômica comprometida deve, é claro, corresponder a uma área do cérebro envolvida no controle da atividade motora.

A análise conjunta dos resultados das duas avaliações, funcional e estrutural, comparando áreas subativadas e superativadas com áreas de maior ou menor volume neuronal, revelou que apenas uma delas atendia a todos os requisitos elencados acima: o núcleo caudado. O núcleo caudado, estrutura subcortical tradicionalmente associada ao planejamento da atividade motora, mostrou uma diminuição significativa de volume bilateral

⁶ Esta última predição fundamenta-se no conhecimento corrente acerca da plasticidade cerebral na criança: se a área “lesada” fosse unilateral, tipicamente esperaria-se que o cérebro se reorganizasse, repassando a responsabilidade pela função comprometida à área homóloga no hemisfério contralateral preservado.

na análise volumétrica⁷ e, na análise funcional, mostrou superativação no hemisfério esquerdo – tornando-se, a partir daí, um grande foco de atenções na investigação do déficit dos KE.

Os estudos de que se seguiram a partir de então reafirmaram a relação do núcleo caudado, estrutura integrante dos núcleos da base, com o déficit de linguagem da família KE. Watkins et al. (2002b) demonstraram que a redução de volume de ambos os núcleos caudados é de aproximadamente 25% nos membros afetados, comparados aos não-afetados e a controles normais pareados por idade, e, de modo ainda mais significativo, observaram uma correlação entre o volume do núcleo caudado e o desempenho em um teste de praxia oral e no teste de repetição de pseudopalavras. Assim, indivíduos com menor volume de caudado tinham pior desempenho em tarefas lingüísticas e não-lingüísticas, e um aumento do volume mostrou-se associado a melhores escores. Esses resultados sugerem que o desenvolvimento anormal do núcleo caudado pode estar na origem de ambos os déficits, motor e cognitivo, observados nos KE.

A avaliação do padrão de atividade cerebral em uma tarefa de morfologia derivacional também resultou na identificação dos núcleos da base como estruturas diferencialmente ativadas nos membros afetados da família KE (Liégeois et al., 2003). A produção de um verbo a partir de um substantivo apresentado pelo entrevistador foi avaliada em duas etapas: em um primeiro momento, com o objetivo de analisar apenas o acesso ao léxico, os indivíduos deveriam apenas “pensar” a resposta, sem dizê-la em voz alta, o que deveria acontecer apenas numa segunda etapa. Foram feitas ainda imagens da atividade cerebral em uma tarefa de repetição de palavras. Enquanto os membros não-afetados da família mostraram uma ativação principalmente do hemisfério esquerdo, envolvendo a área de Broca, na produção de verbos e uma ativação mais bilateral na tarefa de repetição, os membros afetados

⁷ A análise volumétrica consistiu em uma morfometria baseada em voxel (VBM), técnica computacional desenvolvida para comparar e identificar diferenças regionais sutis entre grupos de imagens de tecido cerebral, através de uma análise voxel por voxel (voxel é a unidade tridimensional básica das imagens cerebrais).

apresentaram um padrão mais difuso de ativação bilateral em ambas as tarefas, sugerindo possivelmente uma não-dominância do hemisfério esquerdo para a linguagem.

Tanto nas tarefas silenciosas quanto nas orais, os membros afetados apresentaram um padrão significativo de subativação em diversas áreas corticais e subcorticais, mas de forma mais marcante em duas regiões: a área de Broca (subativada também na região homóloga direita) e o putâmen. Por fim, não foi observada diferença entre os grupos quanto ao tempo para a geração da resposta, a partir do substantivo-estímulo, sugerindo que não há um déficit de acesso ao léxico entre os membros afetados.

A subativação da área de Broca incita uma comparação dos membros afetados aos indivíduos com afasia expressiva, comparação esta reforçada pela observação de Watkins et al. (2002a) de que há marcada semelhança entre os déficits lingüísticos dos dois grupos. Ambos apresentam dificuldade particularmente em tarefas que exigem a repetição de pseudopalavras, a compreensão de estruturas sintáticas complexas e o uso de habilidades morfológicas flexionais e derivacionais. Contudo, os dois grupos diferem fortemente no padrão de acesso lexical, medido pelo tempo de latência para a geração de respostas com a extensão de uma única palavra (como em tarefas de nomeação de objetos, p.ex.). Como reafirma o estudo de Liégeois et al. (2003), os membros afetados da família não apresentam aumento desse tempo de latência e, relatam os autores, também na conversa natural, ao contrário dos afásicos, os membros afetados da família KE não parecem ter dificuldade para “encontrar” as palavras.

A associação dos núcleos da base ao distúrbio observado na família KE foi reforçada pela descoberta de que essa região é um dos locais de expressão do gene *FOXP2* durante o desenvolvimento cerebral normal. O primeiro sinal de expressão do gene foi observado por volta de 45 dias de gestação, no encéfalo posterior (para detalhes, ver Lai et al., 2003); por volta de dois meses de gestação, quando o embrião assume status de feto, o padrão de

expressão do *FOXP2* já se mostra mais difuso, sendo agora identificado no cerebelo, tálamo, núcleo caudado e bulbo – à exceção deste, todos os demais são estruturas predominantemente associadas ao controle motor. O estudo foi também conduzido em ratos, que mostraram um padrão bastante semelhante de expressão temporal e espacial, sugerindo que o *Foxp2* possa estar relacionado ao desenvolvimento de circuitos motores nos mamíferos.

Como o padrão de expressão do *FOXP2* não se revelou nem difuso nem restrito, sendo observado em diferentes áreas, talvez um achado mais relevante seja o das regiões onde ele não foi identificado em nenhum estágio do desenvolvimento humano. Dentre essas, os autores destacam o hipocampo – estrutura tradicionalmente associada à memória declarativa (ou explícita), ou seja, a memória para fatos (conhecimentos) e eventos (episódios), incluindo a memória semântica ou lexical. As memórias declarativas, como o conhecimento dos itens que formam o léxico de uma língua, são construídas tipicamente através de processos associativos, ou seja, através da associação de um estímulo (visual ou tátil, p.ex.) a outro estímulo (sonoro, p.ex). Assim, um estímulo novo provoca a “reação do ‘Que é isto?’”, que compreende um certo grau de alerta e direcionamento da cabeça e dos olhos em direção à fonte de estímulo, ativando respostas exploratórias, como se observa no processo de aquisição lexical. A repetição do estímulo leva à habituação, ou seja, à supressão gradual dessa reação de alerta (Izquierdo, 2002).

A observação de que, durante a embriogênese, o *FOXP2* não se expressa em estruturas hipocâmpais sugere que esse gene não seja importante para o desenvolvimento dessa região do cérebro e, conseqüentemente, as funções a ela relacionadas devem estar preservadas nos indivíduos afetados pela mutação. De fato, é isso que mostram os resultados aqui apresentados. Tanto Gopnik quanto Vargha-Khadem e seus respectivos colaboradores relatam que, além do déficit motor (associado principalmente ao córtex motor e núcleos da base), o déficit lingüístico está presente no domínio da morfossintaxe (associado ao córtex frontal

esquerdo), com preservação das atividades de nomeação e acesso lexical (associadas ao hipocampo).

Neste ponto, é possível avançar uma hipótese explanatória para esse conjunto de achados. Como discutido anteriormente, o distúrbio observado entre parte dos membros da família KE foi prontamente classificado como um tipo particular de *specific language impairment* (SLI), uma vez que os relatos iniciais de Hurst et al. (1990) e Gopnik & Crago (1991) apontavam para uma dispraxia verbal, afetando a articulação dos sons da fala, e para um dano seletivo na aplicação de regras morfossintáticas. Daí as fantasiosas especulações iniciais acerca da possível identificação de um “gene da linguagem”, relatadas por Pinker (2002). Contudo, a investigação subsequente dessa família identificou claramente a existência concomitante de um déficit oromotor, que afetava principalmente, mas não apenas, a articulação dos sons da fala. A relação entre o déficit lingüístico e o déficit motor (se há mesmo alguma) passou, então, a ser um novo objeto de investigação, potencialmente valioso para o entendimento da arquitetura da faculdade da linguagem, especialmente no que diz respeito às alegações de independência desse sistema em relação aos demais “módulos” da cognição (cf. Fodor, 1990).

A etiologia genética do SLI foi sugerida por um número significativo de estudos, assim como a sua relativa herdabilidade (Newbury et al., 2005). Contudo, foram poucos os avanços na identificação das bases moleculares do distúrbio, principalmente em razão da heterogeneidade da classificação. Nos últimos anos, vem-se acumulando um conjunto crescente de evidências de que o SLI, apesar do nome, não compromete apenas a linguagem, mas também funções não-lingüísticas, como habilidades motoras e memória de trabalho (Leonard, 1998; Hill, 2001).

Assim como no caso específico da família KE, as teorias desenvolvidas para explicar o SLI, em seus vários subtipos, oscilaram entre dois paradigmas: déficit gramatical seletivo ou

déficit cognitivo geral, afetando memória e processamento. Em linhas gerais, um sustentava que os indivíduos com SLI (alguns subgrupos, pelo menos) apresentavam um déficit ou um atraso seletivo no desenvolvimento da gramática, i.e., na aplicação de regras para formação de estruturas complexas, enquanto o outro defendia que o SLI decorria de limitações cognitivas que interferiam na capacidade de memória e processamento. Naturalmente, cada uma dessas explicações foi capaz de justificar uma parte dos dados obtidos em grupos de crianças com SLI, como, por exemplo, a não utilização de verbos auxiliares e de formas marcadas de passado ou plural, de um lado, e, do outro, as dificuldades no processamento de estímulos verbais e não-verbais rápidos e na discriminação fonológica (Ullman & Pierpoint, 2005).

A aparente heterogeneidade dos déficits cognitivo e motor, lingüístico e não-lingüístico, observados nos diversos subgrupos de SLI – incluindo a família KE – esconde, contudo, algumas semelhanças importantes. As habilidades gramaticais, que estão comprometidas nesses indivíduos, – ao contrário das habilidades lexicais – seguem regras computacionais próprias, capazes de gerar estruturas compostas, de crescente complexidade. O sistema de regras rege a organização sintática e morfológica e, uma vez aprendido, é aplicado em transformações plenamente produtivas, ou seja, naquelas que podem ser aplicadas a um número potencialmente infinito de novos itens de estrutura semelhante, fundamentalmente, em operações que envolvam o acréscimo de um novo item, como os procedimentos de afixação. As regras computacionais envolvem, assim, a manipulação de símbolos e especificam relações hierárquicas e de encadeamento entre formas morfológicas, como nos seguintes processos: “*look + -ed → looked*”; “*book + -s → books*”; “*happy + -ness → happiness*”; “*book + -shelf → bookshelf*”.

De acordo com o modelo de organização da linguagem em sistema computacional e léxico, e amparando o funcionamento da linguagem nesse sistema dual (seguindo o modelo de *Words and Rules* de Pinker, 1991, 2002), a aquisição, o armazenamento e o acesso a esse tipo

de regra são distintos daqueles usados pelo sistema lexical. O léxico mental está baseado em um procedimento de memória associativa, no qual as representações fonológicas e os conceitos são mapeados uns aos outros, permitindo a aprendizagem, o armazenamento e o processamento dessas relações. Como a relação entre forma e significado nas línguas é arbitrária, não há como prever ou antecipar, simplesmente com base na forma, o significado da palavra “gato”. O fato de que “gato” refere-se a um tipo de animal doméstico peludo de quatro patas precisa ser memorizado, e o conjunto de pares forma-significado forma o léxico mental. Ao contrário, as regras gramaticais, aplicadas aos itens lexicais, são procedimentos composicionais que permitem antecipar o significado a partir das partes, como em “gato + -s” (regra essa que pode, então, ser reaplicada a inúmeras instâncias). O léxico é, assim, uma subdivisão da memória contendo, entre outras coisas, milhares de pares arbitrários de som e significado; a gramática, por sua vez, é um sistema de operações combinatórias produtivas, que fusionam morfemas e palavras do léxico em sintagmas e frases complexas (Pinker & Ullman, 2002).

Assim, enquanto a aquisição e o funcionamento do léxico baseiam-se no estabelecimento de relações arbitrárias – em um sistema de memória declarativa –, a aquisição e o funcionamento da gramática baseiam-se no estabelecimento e na aplicação de regras produtivas – em um sistema de memória procedural. A memória declarativa, através de procedimentos associativos, aprende, armazena e processa todo o léxico mental, além de todo o conhecimento não-lingüístico sobre fatos e eventos; ou seja, ela não apenas informa o significado da palavra *gato*, mas também o nome do gato do vizinho, por exemplo. Ela é explícita, no sentido de que é possível falar sobre ela e até eventualmente lembrar-se como ela foi aprendida. O sistema de regras, por outro lado, encarrega-se do aprendizado implícito e da execução inconsciente não somente de regras morfossintáticas, mas também de atividades cognitivas e motoras, tais como andar de bicicleta, jogar tênis ou dirigir um carro.

Tal proposta está baseada no modelo de Ullman (1999, 2001a, 2001b, 2005), conhecido como Modelo Declarativo/Procedural do Léxico e da Gramática, o qual propõe que a distinção entre léxico e gramática na língua está correlacionada à distinção entre memória declarativa e memória procedural; a primeira está associada à aprendizagem de relações arbitrárias e depende de estruturas do lobo temporal (especialmente o hipocampo); a segunda é especializada na aquisição e na expressão de habilidades motoras e cognitivas inconscientes e depende de circuitos envolvendo o lobo frontal e os núcleos da base.

Na proposta de Ullman, contrastando com modelos anteriores, nenhum dos dois sistemas de memória é exclusivo para a linguagem. Um dos elementos importantes na validação deste modelo é ampla disponibilidade de evidências na literatura sobre correlatos neurais específicos para esses dois subsistemas (ver Ullman & Pierpoint, 2005). A memória associativa envolve os lobos temporais, principalmente o hipocampo. Diversas estruturas cerebrais inter-relacionadas fazem parte do sistema de memória procedural, destacando-se os núcleos da base (especialmente o núcleo caudado) e o cerebelo.

Como afirmado anteriormente, os núcleos da base – caudado, putâmen e globo pálido – são estruturas subcorticais, que estabelecem fortes conexões com o córtex, mais intensamente com os lobos frontais. Essas vias, desde os núcleos da base até o córtex, estão topograficamente organizadas; elas dividem-se em vários circuitos que se projetam para áreas específicas do córtex, e cada circuito subjaz às funções associadas à região cortical no qual se projeta (o circuito que se projeta sobre o córtex motor subjaz a funções motoras, p.ex.). De modo geral, circuitos que passam através do putâmen são importantes para o movimento, enquanto circuitos que passam através do caudado são importantes para aspectos da cognição, muito embora se acredite que ambos sejam importantes em alguma medida para ambas as funções.

Nesse sentido, o circuito que se projeta desde os núcleos da base até a área de Broca e regiões adjacentes do córtex frontal forma a base do sistema procedural envolvido em funções de linguagem, sendo responsável pela aplicação de regras morfosintáticas da gramática mental. A aquisição e o uso dessas regras e operações da gramática mental (como o concatenamento entre verbo e sufixo marcador de passado, “*V + -ed*”) são geralmente implícitos, inconscientes e automatizados.

A memória procedural está relacionada à aprendizagem de novos e ao controle de velhos procedimentos e habilidades. Na definição de Izquierdo (2002),

denominam-se memórias procedurais ou memórias de procedimentos as memórias de capacidades ou habilidades motoras ou sensoriais e o que habitualmente chamamos de “hábitos”. Exemplos típicos são as memórias de como andar de bicicleta, nadar, saltar, soletrar, etc. É difícil “declarar” que possuímos tais memórias; para demonstrar que as temos, devemos de fato andar de bicicleta, nadar, saltar ou soletrar (Izquierdo, 2002:23).

A aquisição desse tipo de procedimento é gradual, acontecendo através de múltiplas experiências e tentativas, mas, uma vez que o procedimento é adquirido, sua aplicação é rápida e automatizada. Por isso, é também chamada de memória implícita, já que o aprendizado e a memória do procedimento não estão disponíveis para serem acessados conscientemente.

A hipótese de um déficit procedural, baseada no Modelo Declarativo/Procedural do Léxico e da Gramática, sustenta que os casos de SLI – e da família KE – podem ser explicados por anormalidades de um único sistema neural de memória, o sistema procedural, que subjaz ao funcionamento da gramática e também ao de outros processos seqüenciais automatizados cujo funcionamento está baseado na atividade de neurônios situados principalmente nos núcleos da base, que se projetam sobre o córtex frontal.

Assim, essa hipótese, ao propor que nos casos de SLI existe um déficit que atinge principalmente o sistema procedural, estabelece que anormalidades em estruturas cerebrais subjacentes a esse sistema devam coexistir com déficits em habilidades motoras e cognitivas. De fato, tais predições (alterações estruturais/funcionais, lingüísticas e não-lingüísticas) foram observadas em subgrupos de portadores de SLI e, de modo importante, na família KE.

Nos membros afetados da família KE, como descrito neste Capítulo, foram identificadas alterações morfofuncionais sobretudo no caudado e no putâmen, estruturas centrais à memória procedural e que demonstraram também corresponder aos locais de grande expressão do gene *FOXP2* durante o processo de desenvolvimento embrionário do sistema nervoso central. No que tange às competências lingüísticas, foi observado um déficit gramatical, aparentemente seletivo, evidenciado por dificuldades na morfologia flexional e derivacional, principalmente na aplicação das regras de flexão verbal. Por fim, em relação às habilidades não-lingüísticas, evidenciou-se um grave déficit na praxia oromotora, afetando de forma marcada, mas não exclusiva, a articulação dos sons da fala.

Os estudos aqui apresentados sugerem, assim, que o *FOXP2* é um gene importante para a formação e o desenvolvimento de determinadas estruturas ou redes neurais, sendo expresso principalmente em regiões associadas ao controle motor e, aparentemente, não sendo expresso em regiões temporais, especialmente no hipocampo, região primordialmente responsável pelas memórias declarativas. Em conformidade com essas observações, os membros afetados da família KE parecem ter preservadas as habilidades lexicais, manifestando um déficit predominantemente oromotor, com dificuldade no rápido seqüenciamento necessário à articulação da fala, mas com evidências também significativas de um déficit morfossintático. Assim, os efeitos da mutação do *FOXP2* parecem, de fato, afetar o processo de aquisição e o funcionamento da linguagem, sem, contudo, comprometer estruturas dedicadas exclusivamente a essa finalidade.

CONCLUSÃO

Na última década vem-se observando um aumento crescente do interesse de pesquisadores de diversas áreas sobre o desenvolvimento da linguagem em contextos patológicos, incluindo síndrome de Down, síndrome de Williams, autismo, dislexia e SLI. O estudo da família KE veio somar-se a esses empreendimentos na tentativa de ampliar o conhecimento acerca da natureza dos déficits lingüísticos e testar teorias sobre as bases cognitivas e biológicas da linguagem.

Dentre as principais questões evocadas por esse tipo de estudo está a possibilidade da observação de déficits seletivos de subcomponentes da linguagem, como aspectos lexicais ou morfossintáticos, assim como a possível identificação de um déficit seletivo da própria linguagem em relação a outros sistemas cognitivos. Tais discussões não são novas; a novidade está no emprego de metodologias alternativas, as quais vêm permitindo uma ampliação do debate. Assim como o estudo, décadas atrás, de lesões focais em áreas encefálicas específicas ajudou a construir o conhecimento atual sobre as relações cérebro-linguagem, hoje o estudo dos transtornos do desenvolvimento da linguagem está amparando a construção e a remodelação de um corpo de evidências ainda mais solidificado sobre tais relações.

Desde as descrições iniciais do SLI, há mais de 20 anos, uma gama de achados levantados a partir da aplicação de testes padronizados e não-padronizados vem sugerindo que essas crianças apresentam dificuldades com o desenvolvimento gramatical, principalmente com aspectos morfológicos. A descrição de Gopnik (1990, 1991) para a família KE está de acordo com essa observação, evidenciando déficits mais marcantes na aplicação e, em menor escala, na compreensão de aspectos morfológicos do inglês, como as marcações de traço para tempo, número e pessoa.

Nessa perspectiva, o SLI representa uma condição na qual a construção de princípios gramaticais é o principal elemento comprometido. Tais limitações na construção das representações gramaticais poderiam resultar de uma experiência restrita com as formas gramaticais da língua, mas este não parece ser o caso da família KE. Alternativamente, pode-se postular que as próprias representações gramaticais seriam, neste caso, deficitárias, devido a uma limitação de base do sistema representacional dedicado à aquisição desses princípios.

Contrastando com essa leitura do SLI como um déficit de competência gramatical, estão as abordagens que tomam em consideração, na análise do distúrbio, fatores externos ao sistema de representação simbólico da gramática e até mesmo fatores externos à faculdade da linguagem. Essas abordagens dão destaque às evidências em favor de um déficit mais difuso do sistema cognitivo, com conseqüências importantes, mas não limitadas, ao processo de aquisição da linguagem. Esta é a posição de um dos principais (se não o principal) pesquisadores do SLI, que vem atualmente se dedicando a mostrar que, embora os aspectos gramaticais sejam um elemento de dificuldade para essas crianças, na comparação entre as diferentes línguas (como inglês, hebreu, italiano, sueco e espanhol), não existe um traço gramatical que uniformemente imponha dificuldades, havendo, contudo, uma dificuldade comum, p.ex., na percepção de formas acusticamente não-salientes (Leonard, 1998).

Não são raros os relatos da identificação de crianças com déficits seletivamente gramaticais; nesses casos a conclusão parece invariavelmente especular sobre tratar-se de uma nova evidência em favor de uma base biológica específica para tais habilidades. Contudo é importante observar que mesmo dados de estudos que não amparam essa separação irão, em muitos casos, apresentar algumas crianças que, de fato, mostram uma aparente dissociação. Em outras palavras, a análise de casos isolados não é suficiente para generalizar conclusões sobre a dissociação da faculdade da linguagem, seja em si própria seja em relação ao sistema cognitivo.

Em relação ao caso único da família KE, seria possível extrapolar as conseqüências da mutação do *FOXP2* nesses indivíduos para uma definição mais ampla do papel desse gene no desenvolvimento da faculdade da linguagem em populações normais? Existem evidências de que sim. Um primeiro argumento em favor da importância do papel do *FOXP2* vem do próprio caso CS, no qual uma translocação balanceada envolvendo o cromossomo 7 (que inclusive acabou facilitando a localização específica do gene) produziu um quadro clínico semelhante ao dos KE afetados.

Contudo, logo após o anúncio da descoberta do *FOXP2*, vários empreendimentos foram lançados com o objetivo de investigar se outras crianças com problemas de linguagem apresentavam também alterações na região 7q31 ou mutações no gene *FOXP2*. Apesar dos esforços, nenhum estudo foi capaz de estabelecer uma associação clara entre uma alteração fenotípica como SLI, autismo ou dislexia e uma variação na seqüência codificadora do *FOXP2* (Newbury et al., 2002; SLI Consortium 2002; O'Brien et al., 2003). Uma das explicações para tal insucesso poderia estar na má caracterização fenotípica: apesar de presente em alguns subgrupos estudados, a dispraxia oral, considerada nuclear no quadro dos KE, não é um déficit proeminente no SLI ou nas demais patologias investigadas.

Seguindo essa lógica, o grupo responsável pela descoberta do *FOXP2* investigou recentemente a seqüência do gene exclusivamente em indivíduos com quadros de dispraxia oral. Em publicação recente do grupo de Vargha-Khadem, MacDermot et al. (2005) divulgaram ter encontrado uma outra mutação no gene *FOXP2* – desta vez a troca de nucleotídeos acarreta uma interrupção na síntese da proteína. Juntamente ao achado molecular, evidenciou-se que, além do indivíduo investigado, sua mãe e sua irmã também possuíam tanto a mutação quanto o déficit de fala e linguagem. Ao demonstrarem um segundo caso de co-segregação familiar de um déficit associado a uma alteração genética específica do *FOXP2*, os pesquisadores ao mesmo tempo reforçam a importância do papel desse gene no

processo de desenvolvimento das capacidades de fala e linguagem e abrem novas possibilidades investigativas, destacando a necessidade de um refinamento dos aspectos fenotípicos a serem procurados nos indivíduos selecionados como “casos” em novos estudos.

No momento, a interpretação mais segura que pode ser feita dessas evidências é a de que o *FOXP2* é um dos muitos genes que afetam sistemas ou mecanismos neurais subjacentes à aquisição e ao funcionamento da linguagem – mas que também subjazem a outras funções. Até agora, as descobertas feitas em torno desse gene específico parecem dar suporte a uma concepção não-unitária de arquitetura da mente, favorecendo uma leitura modular; contudo, a interpretação dos achados específicos aqui analisados sugere a possibilidade de que esses módulos sejam de natureza e organização distintas daquelas previamente propostas na tentativa de circunscrever domínios cognitivos independentes uns dos outros, com bases genéticas próprias.

Esses achados evidenciam ainda que, da identificação de um gene até a definição de sua participação na composição de um fenótipo, há um longo caminho a ser percorrido, que envolve desde a descoberta da alteração molecular específica e a identificação da proteína codificada, passando pelos locais de expressão do gene, até a compreensão da função daquela proteína no desenvolvimento e na atividade de um determinado órgão.

O estudo das relações entre genética e linguagem, como mostra a análise particular da mutação do *FOXP2*, constitui-se em uma nova perspectiva para a exploração de questões fundamentais acerca do desenvolvimento lingüístico e cognitivo. As posições divergentes encontradas entre os pesquisadores não devem causar nem surpresa nem desânimo, dada a complexidade dos problemas enfocados e a relativa novidade dos estudos, os quais só se tornaram possíveis em razão da admirável colaboração dos membros da família KE. Espera-se, assim, que a ampliação do debate traga benefícios não só ao estudo da linguagem, mas também às crianças e às famílias que amparam esta discussão.

Referências bibliográficas

- Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. *Molecular biology of the cell*. 4ª ed. Nova York: Taylor & Francis Group, 2002.
- Alcock KJ, Passingham RE, Watkins KE, Vargha-Khadem. Oral Dyspraxia in Inherited Speech and Language Impairment and Acquired Dysphasia. *Brain Lang* 2000; 75:17-33.
- Alegre M, Gordon P. Frequency effects and the representational status of regular inflections. *J Mem Lang* 1999; 40:41-61.
- Alegre M, Gordon P. Rule-Based versus Associative Processes in Derivational Morphology. *Brain Lang* 1999; 68:347-54.
- American Psychiatric Association (APA). *Manual diagnóstico e estatístico de transtornos mentais – DSM-IV*. 4ª ed. Porto Alegre: Artes Médicas, 1995.
- Anderson SR, Lightfoot DW. The human language faculty as an organ. *Ann Rev Physiol* 1999 (62):697-722.
- Baddeley A. Working memory and language: an overview. *Commun Disord* 2003; 36(3):189-208.
- Baddeley AD, Gathercole SE, Papagno C. The phonological loop as a language learning device. *Psychol Rev* 1998; 105:158-73.
- Bates E. Language and the infant brain. *J Comm Disord* 1999; 32:195-205.
- Bears MF, Connors BW, Paradiso MA. *Neurociências: desvendando o sistema nervoso*. 2ª ed. Porto Alegre: Artes Médicas, 2002.
- Bellugi U, Bihle A, Jernigan T, Trauner D, Doherty S. Neuropsychological, neurological and neuroanatomical profile of Williams syndrome. *Am J Med Genet* 1990; 6:115-125.
- Bellugi U, Lichtenberger L, Mills D, Galaburda A, Korenberg JR. Bridging cognition, the brain and molecular genetics: evidence from Williams syndrome. *Trends Neurosci* 1999; 22(5):197-207.
- Bishop DV, North T, Donlan C. Genetic basis of specific language impairment: evidence from a twin study. *Dev Med Child Neurol* 1995; 37(1):56-71.
- Bishop DV, Snowling MJ. Developmental dyslexia and Specific Language Impairment: Same or different? *Psychol Bull* 2004; 130:858-86.
- Bloom P. Grammatical continuity in language development: the case of subjectless sentences. *Linguist Inq* 1993; 24(4):721-34.
- Bloom P. Subjectless sentences in child language. *Linguist Inq* 1990; 21:491-504.
- Britten RJ. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. *Proc Nat Acad Sci* 2002; 99(21):13633-5.

- Caplan D, Chomsky N. Linguistic Perspectives on Language Development. In: *Biological studies of mental processes*. David Caplan (ed). Cambridge: MIT Press, 1980. p.97-105.
- Carlsson P, Mahlapuu M. Forkhead transcription factors: key players in development and metabolism. *Dev Biol* 2002; 250(1):1-23.
- Chomsky N. A Review of BF Skinner's Verbal Behavior. *Language* 1959; 35(1):26-58.
- Chomsky N. *Aspects of the theory of syntax*. Cambridge: MIT Press, 1965.
- Chomsky N. *Lectures on government and binding*. Dordrecht: Foris, 1981.
- Chomsky N. *Linguagem e Mente*. Brasília: Universidade de Brasília, 1998.
- Chomsky N. Novos Horizontes no Estudo da Linguagem. *DELTA* 1997;13:51-74.
- Chomsky N. *Reflections on language*. Nova York: Pantheon Books, 1975.
- Chomsky N. *Syntactic Structures*. The Hague, Mouton, 1957.
- Chomsky N. *The minimalist program*. Cambridge: MIT Press, 1995.
- Clahsen H & Almazan M. Compounding and inflection in language impairment: Evidence from Williams Syndrome (and SLI). *Lingua* 2001; 110:729-57.
- Clahsen H, Almazan M. Syntax and morphology in Williams syndrome. *Cognition* 1998; 68(3):167-98.
- Clahsen H, Bartke S, Göllner S. Formal features in impaired grammars: a comparison of english and german SLI children. *J Neuroling* 1997; 10:151-71.
- Clahsen H, Dalalakis J. Tense and agreement in Greek SLI: A case study. *Essex Research Reports in Linguistics* 1999; 24: 1-25. Disponível em: <http://privatewww.essex.ac.uk/~harald/papers.htm>.
- Couto HH. *Introdução ao Estudo das Línguas Crioulas e Pidgins*. Brasília: Editora Universidade de Brasília, 1996.
- Elman JL, Bates E, Johnson M, Karmiloff-Smith A, Parisi D, Plunkett K. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge: MIT Press, 1996.
- Fisher SE, Lai CSL, Monaco AP. Deciphering the genetic basis of speech and language disorders. *Annu Rev Neurosci* 2003; 26:57-80.
- Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME. Localization of a Gene Implicated in a Severe Speech and Language Disorder. *Nat Genet* 1998; 18(2): 168-70.
- Fodor JA. *The modularity of mind*. Cambridge: MIT Press, 1990.
- Geschwind N. The organization of language and the brain. *Science* 1970; 170, 940-4.

- Gopnik M, Crago MB. Familial aggregation of a developmental language disorder. *Cognition* 1991; 39:1-50.
- Gopnik M. Feature-blind grammar and dysphasia. *Nature* 1990; 344(6268):715.
- Guttmacher AG, Collins FS. Genomic Medicine - A Primer. *New Engl J Med* 2002; 19:1512-20.
- Haegeman L. *Introduction to Government and Binding theory*. Oxford: Basil Blackwell, 1991.
- Hauser MD, Chomsky N, Fitch WT. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science* 2002; 298:1569-79.
- Hill EL. Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *Int J Lang Commun Disord* 2001; 36(2):149-71.
- Hurst JA, Baraitser M, Auger E, Graham F, Norell S. An extended family with a dominantly inherited speech disorder. *Dev Med Child Neurol* 1990; 32:347-55.
- Hyams N. *Language acquisition and the theory of parameters*. Dordrecht: Reidel, 1996.
- International Human Genome Sequencing Consortium, IHGSC. Initial Sequencing and Analysis of the Human Genome. *Nature* 2001; 409:860-921.
- Izquierdo I. *Memória*. Porto Alegre: Artmed, 2002.
- Jackendoff R. *Foundations of language: brain, meaning, grammar, evolution*. New York: Oxford University Press, 2002.
- Jackendoff R. *The Architecture of the Language Faculty*. Cambridge: MIT Press, 1997.
- Jarrold C, Baddeley AD, Hewes AK. Verbal and nonverbal abilities in the Williams syndrome phenotype: evidence for diverging developmental trajectories. *J Child Psychol Psychiatry* 1998; 39(4):511-23.
- Junqueira LCU, Carneiro J. *Histologia básica*. 9ª ed. Rio de Janeiro: Guanabara Koogan, 1999.
- Kandel ER, Schwartz JH, Jessel TM. *Fundamentos da neurociência e do comportamento*. Rio de Janeiro: Guanabara Koogan, 1997.
- Krause M, Penke M. Inflectional morphology in German Williams syndrome. *Brain Cogn* 2002; 48(2-3):410-3.
- Lai CS, Gerrelli D, Monaco AP, Fisher SE, Copp AJ. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain* 2003; 126(11):2455-62.
- Lai CSL, Fisher SE, Hurst J, et al. The SPCH1 region on human 7q31: genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *Am J Hum Genet* 2000; 67:357-68.

- Lai CSL, Fisher SE, Hurst J, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 2001; 413(6855):519-23.
- Laws G, Bishop DV. Verbal deficits in Down's syndrome and specific language impairment: a comparison. *Int J Lang Commun Disord* 2004; 39(4):423-51.
- Lehmann OJ, Sowden JC, Carlsson P, Jordan T, Bhattacharya SS. Fox's in development and disease. *Trends Genet* 2003; 19:339-44.
- Lenneberg E. *Fundamentos biológicos del lenguaje*. Madrid: Alianza Editorial, 1967.
- Leonard LB. *Children with specific language impairment*. Cambridge: MIT Press, 1998.
- Leonard LB. Language learnability and specific language impairment in children. *Appl Psycholinguist* 1989; 10:179-202.
- Liegeois F, Baldeweg T, Connelly A, Gadian DG, Mishkin M, Vargha-Khadem F. Language fMRI abnormalities associated with *FOXP2* gene mutation. *Nat Neurosci* 2003; 6(11):1230-7.
- Lightfoot D. The child's trigger experience: degree-0 learnability. *Behav Brain Sci* 1989; 12:321-75.
- MacDermot KD, Bonora E, Sykes N. Identification of *FOXP2* truncation as a novel cause of developmental speech and language deficits. *Am J Hum Genet* 2005; 76:1074-80.
- MacWhinney B. *The Childes Project: Tools for Analyzing Talk*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1991.
- Marcus G, Brinkmann U, Clahsen H, Wiese R, Pinker S. German inflection: the exception that proves the rule. *Cognitive Psychol* 1995; 29:189-256.
- Marcus GF, Fisher SE. FOXP2 in focus: what can genes tell us about speech and language? *Trends Cogn Sci* 2003; 7(6):257-62.
- Meisel J. Parâmetros na aquisição. In: Fletcher P, MacWhinney B. *Compêndio da linguagem da criança*. Porto Alegre: Artes Médicas, 1997.
- Menuzzi S. Sobre a evidência para a maturação de universais lingüísticos. *Letras de Hoje* 2001; 36(3):141-66.
- Morange M. *The Misunderstood Gene*. Cambridge: Harvard University Press, 2001.
- Morris CA, Mervis CB. Williams syndrome and related disorders. *Annu Rev Genomics Hum Genet* 2000; 1:461-84.
- Munson B, Kurtz BA, Windsor J. The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *J Speech Lang Hear Res* 2005; 48(5):1033-47.
- Newbury DF, Bishop DV, Monaco AP. Genetic influences on language impairment and phonological short-term memory. *Trends Cogn Sci* 2005; 9(11):528-34.

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. *Am J Hum Genet* 2002; 70(5):1318-27.

Nicolosi L, Harryman E, Kresheck J. *Vocabulário dos distúrbios da comunicação: fala, linguagem e audição*. 3ª ed. Porto Alegre: Artmed, 1996.

O'Brien EK, Zhang X, Nishimura C, Tomblin JB, Murray JC. Association of specific language impairment (SLI) to the region of 7q31. *Am J Hum Genet* 2003; 72:1536-43.

Ojopi EPB, Gregório SP, Guimarães PEM, Fridman C, Dias Neto E. O genoma humano e as perspectivas para o estudo da esquizofrenia. *Rev Psiq Clín* 2004; 31(1):9-18.

Page GP, George V, Go RC, Page PZ, Allison DB. "Are We There Yet?": Deciding When One Has Demonstrated Specific Genetic Causation in Complex Diseases and Quantitative Traits - Invited Editorial. *Am J Hum Genet* 2003; 73:711-9.

Peña M, Maki A, Kovacic D, Dehaene-Lambertz G, Koizumi H, Bouquet F, Mehler J. Sounds and silence: an optical topography study of language recognition at birth. *Proc Natl Acad Sci USA* 2003; 100(20):11702-5.

Pinker S. *O Instinto da Linguagem: como a mente cria a linguagem*. São Paulo: Martins Fontes, 2002.

Pinker S. Rules of language. *Science* 1991; 253:530-535.

Pinker S. Talk of genetics and vice versa. *Nature* 2001; 413:465-7.

Pleth C, Lukacs A, Racsmany M. Morphological patterns in Hungarian children with Williams syndrome and the rule debates. *Brain Lang* 2003; 86(3):377-83.

Raposo E. *Teoria da Gramática: A faculdade da linguagem*. Lisboa: Ed. Caminho, 1992. 527p.

Ridley M. *Nature via nurture: genes, experience, and what makes us human*. Nova York: Harper Collins, 2003.

Rodriguez N. *Neurolingüística dos distúrbios da fala*. São Paulo: Cortez, 1999.

Scarpa EM. Aquisição da linguagem. In: Mussalim F, Bentes AC (org.). *Introdução à lingüística: domínios e fronteiras*. Vol 2. São Paulo: Ed. Cortez, 2000.

Sherman SL, DeFries JC, Gottesman II, et al. Behavioral genetics '97: ASHG statement. Recent developments in human behavioral genetics: past accomplishments and future directions. *Am J Hum Genet* 1997; 60(6):1265-75.

Silva Nunes MVR. Aprendizagem da leitura e o "loop" fonológico. *RFML* 2001;6(1):21-8. Disponível em: http://www.fm.ul.pt/public/ge/Index_GE/2001/1/PDF/21.pdf

Skinner BF. *Verbal behavior*. Nova York: Appleton Century Crofts, 1957.

SLI Consortium. A genomewide scan identifies two novel loci involved in specific language impairment. *Am J Hum Genet* 2002; 70(2):384-98.

Stromswold K. The heritability of language: a review and metaanalysis of twin, adoption and linkage studies. *Language* 2001; 77:647-723.

Tallal P, Hirsch LS, Realpe-Bonilla T, Miller S, Brzustowicz LM, Bartlett C, et al. Familial aggregation in specific language impairment. *J Speech Lang Hear Res* 2001; 44(5):1172-82.

Thompson PM, Cannon TD, Narr KL, et al. Genetic Influences on Brain Structure. *Nature Neurosci* 2001; 4(12):1253-8.

Tomasello M. Language is not an instinct. Review of Pinker's "The language instinct". *Cognitive Dev* 1995; 10:131-56.

Tomblin JB, Buckwalter PR. Heritability of poor language achievement among twins. *J Speech Lang Hear Res* 1998; 41(1):188-99.

Tomblin JB, Pandich J. Lessons from children with specific language impairment. *Trends Cogn Sci* 1999; 3(8):283-5.

Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *J Speech Lang Hear Res* 1997; 40:1245-60.

Tomblin JB, Records NL, Zhang X. A system for the diagnosis of specific language impairment in kindergarten children. *J Speech Hear Res* 1996; 39:1284-94.

Ullman MT, Pierpont EI. Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex* 2005; 41(3):399-433.

Ullman MT. A neurocognitive perspective on language: the declarative/procedural model. *Nat Rev Neurosci* 2001;2(10):717-26.

Ullmann MT. The declarative/procedural model of lexicon and grammar. *J Psycholinguist Res* 2001; 30(1):37-69.

Ullman MT, Gopnik M. Inflectional morphology in a family with inherited specific language impairment. *Applied Psycholinguistics* 1999; 20:51-117.

Vargha-Khadem F, Passingham R. Speech and language defects. *Nature* 1990; 346:226.

Vargha-Khadem F, Watkins K, Alcock K, Fletcher P, Passingham RE. Cognitive and praxic deficits in a large family with a genetically transmitted speech and language disorder. *Proc Natl Acad Sci* 1995; 92:930-3.

Vargha-Khadem F, Watkins KE, Price CJ, Ashburner J, Alcock KJ, Connelly A, et al. Neural basis of an inherited speech and language disorder. *Proc Natl Acad Sci* 1998; 95:12695-700.

Venter CJ, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The Sequence of the Human Genome. *Science* 2001; 291:1304-51.

Vicari S, Bates E, Caselli MC, Pasqualetti P, Gagliardi C, Tonucci F, Volterra V. Neuropsychological profile of Italians with Williams syndrome: an example of a dissociation between language and cognition? *J Int Neuropsychol Soc* 2004; 10(6):862-76.

Watkins KE, Dronkers NF, Vargha-Khadem F. Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain* 2002; 125:452-64.

Watkins KE, Vargha-Khadem F, Ashburner J, Passingham RE, Friston KJ, Connelly A, Frackowiak RSJ, Mishkin M, Gadian DG. MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain* 2002; 125: 465-78.

Watson JD, Crick FHC. Molecular Structure in Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 1953; 171:737-8.

Weedwood B. *História concisa da lingüística*. São Paulo: Parábola Editorial, 2002.

Wexler K. Innateness and maturation in linguistic development. *Dev Psychobiol* 1990; 23(7):645-60.

Ziegler JC, Pech-Georgel C, George F, Alario F-X, Lorenzi C. Deficits in speech perception predict language learning impairment. *Proc Natl Acad Sci* 2005; 102(39), 14110-5.