

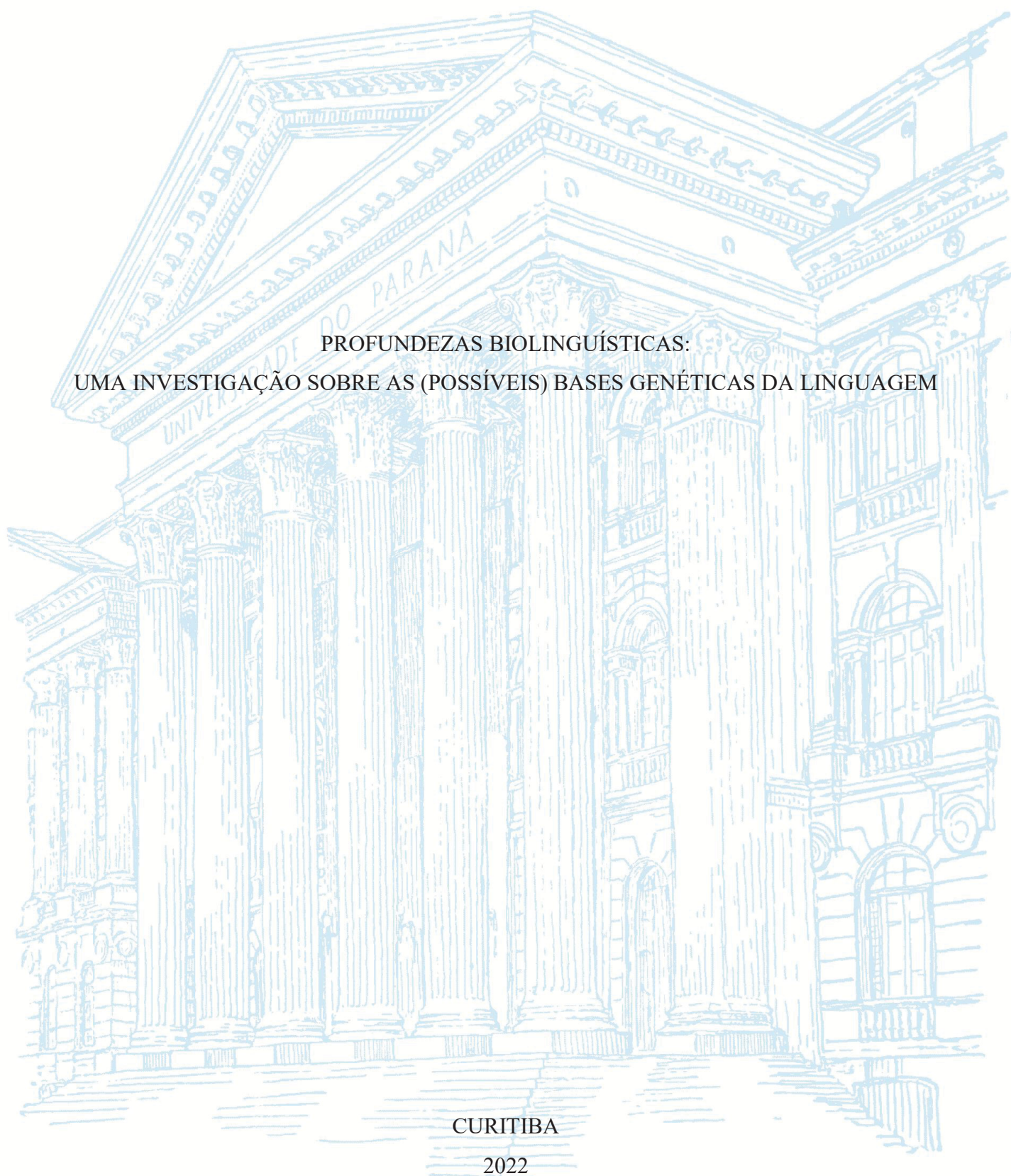
UNIVERSIDADE FEDERAL DO PARANÁ

RODRIGO PEREIRA DOS ANJOS

PROFUNDEZAS BIOLINGÜÍSTICAS:
UMA INVESTIGAÇÃO SOBRE AS (POSSÍVEIS) BASES GENÉTICAS DA LINGUAGEM

CURITIBA

2022



RODRIGO PEREIRA DOS ANJOS

PROFUNDEZAS BIOLINGUÍSTICAS:
UMA INVESTIGAÇÃO SOBRE AS (POSSÍVEIS) BASES GENÉTICAS DA LINGUAGEM

Dissertação apresentada ao curso de Pós-Graduação em Letras, Setor de Ciências Humanas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Letras.

Orientadora: Profa. Dra. Patrícia de Araújo Rodrigues

CURITIBA

2022

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA

Anjos, Rodrigo Pereira dos
Profundezas biolinguísticas : uma investigação sobre as
(possíveis) bases genéticas da linguagem. / Rodrigo Pereira dos
Anjos. – Curitiba, 2022.

1 recurso on-line : PDF.

Mestrado (Dissertação em Letras) – Universidade Federal do
Paraná, Setor de Ciências Humanas, Programa de Pós-
Graduação em Letras.

Orientadora: Prof^a. Dr^a. Patrícia de Araújo Rodrigues

1. Biolinguística. 2. Linguagem e línguas – Aspectos
fisiológicos. 3. Gene. I. Rodrigues, Patrícia de Araújo, 1960-.
II. Universidade Federal do Paraná. Programa de Pós-Graduação
em Letras. III. Título.

Bibliotecária: Fernanda Emanoéla Nogueira Dias CRB-9/1607

ATA Nº1111

ATA DE SESSÃO PÚBLICA DE DEFESA DE MESTRADO PARA A OBTENÇÃO DO GRAU DE MESTRE EM LETRAS

No dia cinco de abril de dois mil e vinte e dois às 14:00 horas, na sala 1020, Ed. D. Pedro I, foram instaladas as atividades pertinentes ao rito de defesa de dissertação do mestrando **RODRIGO PEREIRA DOS ANJOS**, intitulada: **PROFUNDEZAS BIOLINGÜÍSTICAS: Uma investigação sobre as (possíveis) bases genéticas da linguagem**, sob orientação da Profa. Dra. PATRÍCIA DE ARAUJO RODRIGUES. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação LETRAS da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: PATRÍCIA DE ARAUJO RODRIGUES (UNIVERSIDADE FEDERAL DO PARANÁ), MARCUS VINICIUS DA SILVA LUNGUINHO (UNIVERSIDADE DE BRASÍLIA), MARIA CRISTINA FIGUEIREDO SILVA (UNIVERSIDADE FEDERAL DO PARANÁ). A presidência iniciou os ritos definidos pelo Colegiado do Programa e, após exarados os pareceres dos membros do comitê examinador e da respectiva contra argumentação, ocorreu a leitura do parecer final da banca examinadora, que decidiu pela **APROVAÇÃO**. Este resultado deverá ser homologado pelo Colegiado do programa, mediante o atendimento de todas as indicações e correções solicitadas pela banca dentro dos prazos regimentais definidos pelo programa. A outorga de título de mestre está condicionada ao atendimento de todos os requisitos e prazos determinados no regimento do Programa de Pós-Graduação. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, PATRÍCIA DE ARAUJO RODRIGUES, lavrei a presente ata, que vai assinada por mim e pelos demais membros da Comissão Examinadora.

CURITIBA, 05 de Abril de 2022.

Assinatura Eletrônica

11/04/2022 11:01:30.0

PATRÍCIA DE ARAUJO RODRIGUES

Presidente da Banca Examinadora

Assinatura Eletrônica

05/04/2022 16:20:23.0

MARCUS VINICIUS DA SILVA LUNGUINHO

Avaliador Externo (UNIVERSIDADE DE BRASÍLIA)

Assinatura Eletrônica

05/04/2022 21:28:58.0

MARIA CRISTINA FIGUEIREDO SILVA

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação LETRAS da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **RODRIGO PEREIRA DOS ANJOS** intitulada: **PROFUNDEZAS BIOLINGÜÍSTICAS: Uma investigação sobre as (possíveis) bases genéticas da linguagem**, sob orientação da Profa. Dra. PATRÍCIA DE ARAUJO RODRIGUES, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 05 de Abril de 2022.

Assinatura Eletrônica

11/04/2022 11:01:30.0

PATRÍCIA DE ARAUJO RODRIGUES

Presidente da Banca Examinadora

Assinatura Eletrônica

05/04/2022 16:20:23.0

MARCUS VINICIUS DA SILVA LUNGUINHO

Avaliador Externo (UNIVERSIDADE DE BRASÍLIA)

Assinatura Eletrônica

05/04/2022 21:28:58.0

MARIA CRISTINA FIGUEIREDO SILVA

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

AGRADECIMENTOS

Durante um tempo, hesitei em escrever esta seção. Não porque não me sentia grato às pessoas que cito aqui – sou imensamente grato –, mas porque este sempre me pareceu um gênero textual utilizado sobretudo para cumprir formalidades. De fato, eu, como leitor, nunca havia me interessado pela seção de agradecimentos, simplesmente porque todos aqueles nomes citados – aparentemente tão queridos ao autor – não me diziam nada. Mas agora que estou do lado de cá da escrita, compreendi que esta não é uma seção dedicada ao leitor comum – estas são palavras que só adquirem sentido pleno quando lidas pelas pessoas que de alguma forma me auxiliaram nesta caminhada. A cada um de vocês, os meus mais profundos agradecimentos.

À minha professora e orientadora, Patrícia de Araújo Rodrigues, não apenas pela leitura atenta do trabalho e pelas sugestões, mas por ter lançado a semente desta pesquisa numa aula sobre aquisição de linguagem no meu primeiro semestre da graduação. Foi ali que eu entrei em contato pela primeira vez com a hipótese chomskyana de que a linguagem é uma dotação biológica condicionada geneticamente. De certa forma, esta dissertação começou no momento em que eu perguntei se alguém do nosso setor estava estudando a relação entre genes e linguagem e a professora respondeu que não. Passados alguns anos, cá estamos. Agradeço a você também por ter ajudado a despertar em mim o amor pela sintaxe. Muito obrigado.

À professora Maria Cristina Figueiredo Silva e ao professor Marcus Vinicius da Silva Lunguinho, pela leitura criteriosa do texto e pelas mudanças sugeridas na banca de qualificação. Cada apontamento foi fundamental para que eu refinasse o trabalho e repensasse certas decisões – e se eventualmente algumas mudanças não foram adotadas, isto se deve única e exclusivamente à teimosia deste que vos fala. Agradeço também por terem aceitado o convite para participar da banca de defesa. Obrigado.

A todos os professores do departamento de Letras da Universidade Federal do Paraná, especialmente àqueles de quem eu fui aluno, na graduação e na pós. Obrigado por terem participado da minha formação e por terem me inspirado a estudar cada vez mais. Agradeço ao professor Luís Bueno, pelo bom-humor nas aulas e pela orientação da minha monografia, e às professoras Gesualda Lourdes dos Santos Rasia e Teresa Cristina Wachowicz, por terem

me acolhido no Pibid durante a graduação e por terem me dado a oportunidade de começar minha atividade como professor.

Aos meus colegas (presenciais) da graduação e aos meus colegas (virtuais) da pós. Algumas amizades se estenderam por mais tempo, enquanto outras duraram apenas um semestre. Muitas vezes, a nossa convivência com alguém na universidade se resume a um único bate-papo em uma noite fria, no intervalo entre uma aula e outra – e eu tenho alguns desses brevíssimos encontros gravados na memória. Por isso, independentemente da duração e da proximidade, agradeço a cada um de vocês por terem dividido este pedaço de tempo comigo e por fazerem parte da minha vida. Também agradeço aos amores (correspondidos e não correspondidos) desse período. Sem vocês, eu tenho certeza que esses anos na universidade seriam mais tranquilos, mas teriam bem menos graça.

À minha mãe, que sempre me apoiou em todas as minhas decisões e que sempre esteve ao meu lado. Você, sem dúvida, mãe, é o meu porto seguro – o lugar para onde eu sempre volto. Obrigado por tudo.

À Universidade Federal do Paraná e ao Programa de Pós-Graduação em Letras, pela oportunidade de ter acesso a uma educação de ponta.

À Capes, pelo auxílio financeiro.

Por último, gostaria de agradecer não *a* alguém em especial, mas *pela* minha vida, por mais abstrato que isso possa parecer. Ao longo deste texto, eu menciono algumas vezes o quão fascinante – e misterioso – é estar vivo. Quando penso na imensidão do universo e na quantidade incalculável de fenômenos que nos trouxeram até aqui, sinto-me invadido por um sentimento imenso de gratidão por ter a chance única de experimentar a vida. Poder levantar pela manhã e ver o Sol nascendo enquanto tomo uma xícara de café são motivos suficientes para que eu seja eternamente grato por esse período curto de tempo no qual existimos profundamente.

RESUMO

Este trabalho tem como objeto de investigação as (possíveis) relações entre genes e linguagem. Desde que a Gramática Gerativo-Transformacional (GGT) assumiu como uma de suas principais hipóteses de trabalho a ideia chomskyana de que a faculdade da linguagem é uma dotação biológica condicionada geneticamente, assumiu também o compromisso de estabelecer conexões plausíveis entre o programa inato de aquisição de uma língua qualquer (tradicionalmente referido como Gramática Universal) e o programa genético subjacente responsável pela codificação bioquímica destas propriedades linguísticas inatas. No entanto, apesar de a GGT ter assumido a etiologia genética da linguagem como um pressuposto fundamental, as relações concretas entre genes e linguagem permanecem relativamente inexploradas. O objetivo principal deste trabalho é exatamente explorar esta zona de interseção entre a Linguística e a Genética, analisando e problematizando um conjunto selecionado de textos e pesquisas relacionados ao tema. Para isso, partimos da definição do seguinte problema: (i) *quais são e de que modo atuam os genes (ou fatores genéticos) que determinam a faculdade da linguagem?*. Longe de oferecermos uma resposta categórica e completa a essa questão (até certo ponto ainda não disponível no atual estágio do conhecimento humano), apresentamos ao longo do trabalho informações, estudos, análises e reflexões que giram em torno de aspectos cruciais do problema delineado em (i). Inicialmente, realçamos a singularidade biológica da faculdade da linguagem (uma dotação cognitiva exclusiva da espécie humana), comparando-a com as competências comunicativas de outros animais; em seguida, apresentamos argumentos a favor da hipótese do inatismo linguístico, segundo a qual os seres humanos nascem com uma gramática geneticamente programada que lhes possibilita adquirir a língua de seu ambiente; a fim de materializarmos a discussão, apresentamos em um dos capítulos uma série de fatores genéticos que, quando afetados por mutações, dão origem a problemas linguísticos de natureza variada; dado o seu caráter paradigmático, dedicamos um capítulo inteiro ao gene *FOXP2* e à família KE (um agrupamento familiar que possui entre seus integrantes indivíduos que apresentam um grave distúrbio de fala e de linguagem derivado de uma mutação pontual no *FOXP2*); ao final, retomamos alguns pontos do problema (i), contextualizando-o no quadro do Programa Minimalista. Apesar dos avanços tecnológicos de mapeamento entre fenótipo e genótipo e dos fortes indícios relacionados à vinculação entre genes e linguagem, ainda há muito o que se fazer a fim de descobrir quais fatores genéticos determinam a faculdade da linguagem, de que forma eles atuam e por quais propriedades linguísticas eles são responsáveis.

Palavras-chave: Biolinguística. Genes e linguagem. *FOXP2*.

ABSTRACT

This thesis has as its object of investigation the (possible) relationships between genes and language. Since the Transformational-Generative Grammar (TGG) assumed as one of its main working hypotheses the Chomskyan idea that the language faculty is a genetically conditioned biological endowment, it has also assumed the commitment to establish plausible connections between the innate program of acquisition of any language (traditionally referred to as Universal Grammar) and the underlying genetic program responsible for the biochemical encoding of these innate linguistic properties. However, despite the fact that TGG has assumed the genetic etiology of language as a fundamental assumption, the concrete relationships between genes and language remain relatively unexplored. The main objective of this work is precisely to explore this area of intersection between Linguistics and Genetics, analyzing and problematizing a selected set of texts and researches related to the topic. For this, we start from the definition of the following problem: (i) *What are the genes (or genetic factors) that determine the language faculty and how do they act?*. Far from offering a categorical and complete answer to this question (to some extent not yet available in the current stage of human knowledge), we present throughout the work information, studies, analyzes and reflections that revolve around crucial aspects of the problem outlined in (i). Initially, we emphasize the biological uniqueness of the language faculty (a cognitive endowment exclusive to the human species), comparing it with the communicative competences of other animals; then, we present arguments in favor of the hypothesis of linguistic innateness, according to which human beings are born with a genetically programmed grammar that allows them to acquire the language of their environment; in order to materialize the discussion, we present in one of the chapters a series of genetic factors that, when affected by mutations, give rise to linguistic problems of a varied nature; given its paradigmatic character, we dedicate an entire chapter to the *FOXP2* gene and the KE family (a family grouping that has among its members individuals who have a severe speech and language disorder derived from a point mutation in *FOXP2*); at the end, we return to some points of problem (i), contextualizing it within the framework of the Minimalist Program. Despite the technological advances in mapping between phenotype and genotype and the strong evidences related to the link between genes and language, much remains to be done to discover which genetic factors determine the faculty of language, how they act and for which linguistic properties they are responsible.

Keywords: Biolinguistics. Genes and language. *FOXP2*.

SUMÁRIO

Capítulo 1 – Genes e linguagem: a apresentação de um (sub)problema.....	11
Capítulo 2 – A singularidade biológica da faculdade da linguagem.....	18
2.1. O código da vida.....	18
2.2. Comunicação animal.....	28
2.2.1. Abelhas dançarinas.....	28
2.2.2. Papagaios, beija-flores e pássaros canoros.....	35
2.2.3. Primatas sinalizadores.....	49
Resumo do capítulo e considerações finais.....	61
Capítulo 3 – Inato vs. adquirido.....	63
3.1. Natureza vs. cultura.....	63
3.2. Órgão da linguagem vs. mecanismos gerais da inteligência.....	82
3.3. Hipótese inatista forte vs. hipótese inatista fraca.....	90
Resumo do capítulo e considerações finais.....	95
Capítulo 4 – (Alguns) fatores genéticos associados à linguagem.....	98
4.1. Modelos e métodos de investigação.....	98
4.2. Alguns exemplos de erros inatos do metabolismo que afetam a linguagem e outras funções cognitivas.....	107
4.2.1. Gene <i>SLC17A5</i>	108
4.2.2. Gene <i>PAH</i>	109
4.2.3. Genes <i>IDUA</i> e <i>SGSH</i>	110
4.2.4. Gene <i>DHCR7</i>	112
4.2.5. Genes <i>GALT</i> , <i>GALK</i> e <i>GALE</i>	114
4.3. Pausa para um balanço epistemológico e linguístico sobre o que temos visto até aqui.....	116
4.4. Distúrbio(s) específico(s) de linguagem.....	125
Resumo do capítulo e considerações finais.....	146
Capítulo 5 – <i>FOXP2</i>	148
5.1. Breve história de uma descoberta.....	152
5.2. Papéis biológicos do <i>FOXP2</i>	162
5.3. Aspectos filogenéticos do <i>FOXP2</i>	167
5.4. A família KE.....	171
Resumo do capítulo e considerações finais.....	202
Capítulo 6 – Breve consideração final.....	205
Apêndice.....	213
Referências.....	216

CAPÍTULO 1 – GENES E LINGUAGEM: A APRESENTAÇÃO DE UM (SUB)PROBLEMA

Nós costumávamos pensar que o nosso destino estava nas estrelas. Agora sabemos que, em boa parte, o nosso destino está nos genes. (James Watson).¹

A relação entre genes e linguagem vem sendo investigada já há algum tempo, com uma atenção especial voltada para o estudo de diversos tipos de síndromes genéticas ou cromossômicas associadas a déficits linguísticos de natureza e grau variados (como a Síndrome de Down e determinados casos de autismo). No entanto, apesar das fortes evidências de participação de fatores genéticos na determinação da cognição linguística humana, foi apenas em 2001, no mesmo ano em que os trabalhos relatando o sequenciamento quase completo do genoma humano foram publicados na *Nature* (LANDER *et al*, 2001) e na *Science* (VENTER *et al*, 2001), que a comunidade científica anunciou a descoberta daquele que seria o primeiro gene relacionado diretamente com o processo de aquisição e desenvolvimento da linguagem: o *FOXP2* (LAI *et al*, 2001).

A identificação do *FOXP2* realçou a natureza biológica da faculdade da linguagem, estabeleceu um ponto de contato entre a Linguística e a Biologia Molecular – exigindo dessas duas áreas do conhecimento um comprometimento mútuo na tentativa de encontrar explicações adequadas para a relação entre cognição linguística e DNA – e representou um marco importante na consolidação da Biolinguística como programa de investigação científica.

A biolinguística é um importante novo campo interdisciplinar que se propõe a explorar as propriedades básicas da linguagem humana e a investigar como ela amadurece no indivíduo, como ela é usada no pensamento e na comunicação, quais circuitos cerebrais a implementam,

¹ “We used to think our fate was in our stars. Now we know, in large measure, our fate is in our genes”. Citado em JAROFF, 1989: 67. Todas as traduções dos trechos citados em outras línguas são de nossa autoria, salvo em algumas poucas passagens em que optamos por uma versão já traduzida do texto original. Nestes casos, indicamos os respectivos tradutores na seção de Referências.

qual combinação de genes a suporta e como ela surgiu em nossa espécie. (BOECKX, 2011: vii)².

A proposta de explorar os aspectos neurofisiológicos, cognitivos, genéticos e evolutivos da linguagem humana faz da Biolinguística uma área de sobreposição de labirintos a serem investigados e mapeados um no outro: o labirinto da linguagem, enredado no labirinto abstrato da cognição, localizado no labirinto histológico do cérebro, codificado no labirinto bioquímico do genoma humano, derivado do labirinto evolutivo que o moldou ao longo do tempo. No limite, busca-se uma descrição biológica unificada da linguagem, concebendo-a como objeto natural.

É possível pensar o programa de pesquisa da Biolinguística como uma tentativa de responder às duas últimas questões propostas por Chomsky & Lasnik (1993: 508)³ na delimitação dos problemas clássicos que caracterizam o estudo da linguagem humana de uma perspectiva cognitiva/neurológica (na formulação que segue, *João* representa um falante de uma língua qualquer).

- (a) O que o João sabe quando sabe uma língua específica?
- (b) Como o João adquiriu esse conhecimento?
- (c) Como o João usa esse conhecimento?
- (d) Como essas propriedades da mente/cérebro evoluíram na espécie?
- (e) Como essas propriedades são realizadas nos mecanismos do cérebro?

Descobrir como o dispositivo neurobiológico que sustenta o conhecimento gramatical dos seres humanos surgiu e evoluiu ao longo da história do *Homo sapiens sapiens* é o objetivo da diretriz programática (d) (*Problema de Darwin-Wallace*), o que exige a cooptação de conhecimentos advindos de áreas como a Biologia Evolutiva, a Antropologia e a Paleontologia; descrever e explicar quais são os mecanismos eletrofisiológicos e quais são as

² “Biolinguistics is an important new interdisciplinary field that sets out to explore the basic properties of human language and to investigate how it matures in the individual, how it is put to use in thought and communication, what brain circuits implement it, what combination of genes supports it, and how it emerged in our species.”

³ “(a) What does Jones know when he has a particular language?
 (b) How did Jones acquire this knowledge?
 (c) How does Jones put this knowledge to use?
 (d) How did these properties of the mind/ brain evolve in the species?
 (e) How are these properties realized in mechanisms of the brain?”

redes sinápticas que caracterizam esse dispositivo neurológico de aquisição/processamento linguístico – e como a faculdade da linguagem emerge desse emaranhado tecidual altamente organizado formado por bilhões de neurônios – são tarefas que correspondem à diretriz (e) (*Problema de Broca*), o que abre espaço para um diálogo interdisciplinar entre a Linguística e as Neurociências, a Fisiologia e as Ciências Computacionais (como a Inteligência Artificial e a Robótica). Cada um desses problemas se desdobra em uma série de questionamentos (ou subtemas) que orientam linhas de pesquisa dedicadas a investigar aspectos distintos dos mecanismos biológicos que subjazem à linguagem. Uma das questões envolvidas no Problema de Darwin-Wallace, por exemplo, é saber se a faculdade da linguagem evoluiu biologicamente de sistemas cognitivos pré-existentes ou se emergiu como uma capacidade de processamento totalmente inédita no curso de nossa evolução, sendo fruto de alguma inovação biológica derivada de mutações do genoma humano. Já questões que dizem respeito à arquitetura do centro nervoso humano – como a modularidade ou não da mente/cérebro – procuram circunscrever a linguagem à substância biofísica que a realiza neurologicamente e são instâncias do Problema de Broca.

Atravessando os problemas (d) e (e) – em uma zona de interseção entre a Biologia Evolutiva e a Neurologia –, há um (sub)problema de natureza própria, que diz respeito ao programa genético subjacente à faculdade da linguagem e às diretrizes bioquímicas que a condicionam. Uma possível formulação desse (sub)problema – que poderíamos chamar de Problema de Mendel⁴ ou Problema de Watson-Crick-Franklin⁵ – encontra-se em (1):

- (1) Quais são e de que modo atuam os genes (ou fatores genéticos) que determinam a faculdade da linguagem?

Esse problema – que nos remete às bases biomoleculares do labirinto genômico do qual deriva a cognição linguística humana – é o motivo condutor que orienta a discussão do

⁴ Realizando cruzamentos entre ervilhas (*Pisum sativum*) com características distintas e lançando mão de ferramentas matemáticas simples aplicadas à análise das frequências e das proporções observadas nas características dos descendentes, Gregor Johann Mendel (1822-1884) propôs, em 1865, as leis fundamentais da hereditariedade (Primeira e Segunda Leis de Mendel).

⁵ Tendo como peça-chave os experimentos fotográficos realizados pela química britânica Rosalind Elsie Franklin (1920-1958), que mostravam como era o padrão de difração de raios-X apresentado pelo DNA, James Dewey Watson (1928) e Francis Harry Compton Crick (1916-2004) propuseram, em 1953, o modelo da dupla-hélice para a molécula do DNA.

presente trabalho, cujo objetivo é investigar as (possíveis) relações entre genes e linguagem, oferecendo uma visão introdutória sobre o assunto a partir da análise e da problematização de um conjunto selecionado de estudos de Linguística, Genética e Biolinguística que se dedicaram ao tema. Neste sentido, a diretriz fundamental que guia a nossa abordagem é baseada na seguinte via genético-linguística:

(2) genes → cérebro/mente → cognição (linguagem)

Esse é o princípio básico, que deve estar sempre em mente durante a análise das questões abordadas, numa visão integrada das três dimensões (genética, neurológica e linguística) que atravessam o problema (1). Esta tríade de dimensões interconectadas pode nos ajudar a mapear as interações entre fatores genéticos e linguagem na medida em que nos oferece uma matriz epistemológica simples a partir da qual podemos abordar adequadamente um problema de natureza relativamente complexa – um problema que, em última análise, consiste em descobrir de que forma uma função biológica *f* foi e é capaz de derivar, a partir de um conjunto de instruções codificadas no genoma do *Homo sapiens*, uma faculdade cognitiva de natureza única no universo (ao menos, no universo conhecido por nós): a cognição linguística.

Também é importante ressaltar que na maior parte deste trabalho estamos assumindo que cérebro (dimensão física) e mente (dimensão abstrata) são contrapartes de uma mesma entidade biológica. Assim, cada fenômeno linguístico corresponde simultaneamente a um fenômeno mental e a um fenômeno material realizado pela substância eletrofisiológica dos neurônios (uma brevíssima discussão sobre o problema da relação entre cérebro e mente pode ser encontrada mais adiante, na nota 55). Sabemos das limitações teóricas dessa abordagem fisicalista, mas, se estamos assumindo que a faculdade da linguagem é determinada geneticamente, inevitavelmente precisamos caracterizá-la a partir das contrapartes materiais (cerebral e genética) que a sustentam.

Além de ser uma breve introdução sobre o tema deste trabalho, o objetivo deste capítulo inicial é apresentar um panorama geral dos assuntos que serão discutidos ao longo do texto. Cada um dos capítulos subsequentes se destina a analisar determinados aspectos do problema (1), buscando, sempre que possível, uma convergência entre a Genética e a Linguística, orientados pela diretriz que estabelecemos em (2). A fim de guiar a leitura, abrimos cada capítulo com um conjunto compacto de perguntas que antecipam os temas principais que serão desenvolvidos no decorrer das seções.

No Capítulo 2, são apresentados os mecanismos genéticos básicos pelos quais a vida se expressa, com uma atenção especial para o código genético e para o modo como as informações biológicas estão codificadas em um “alfabeto” bioquímico universal que atravessa todo o mundo vivo. Além de apresentar brevemente alguns conceitos de genética que serão retomados em discussões posteriores, o objetivo desse capítulo é destacar a singularidade biológica que caracteriza a faculdade da linguagem e que faz dela uma dotação cognitiva única no reino animal. Por isso, são discutidos contrastivamente alguns aspectos comportamentais dos sistemas de comunicação de outros animais – mais especificamente, de abelhas, aves (como papagaios) e primatas superiores (como chimpanzés). Apesar de compartilhar com outras espécies inúmeras características biológicas “transversais” – o que realça a unidade da vida e a universalidade do programa bioquímico que a codifica –, a espécie humana, em termos ontogenéticos e evolutivos, é o resultado de um arranjo de fatores genéticos do qual emergem características biológicas “singulares” – sendo a capacidade de adquirir e desenvolver uma língua natural talvez a mais idiossincrática delas.

No Capítulo 3, são discutidos e problematizados alguns aspectos teóricos da relação entre genes e linguagem sob a perspectiva do inatismo linguístico. Sem a pretensão de ser um panorama que cubra exaustivamente todos os modelos de análise linguística de caráter inatista, o objetivo do capítulo é apresentar uma fundamentação teórica para a discussão de determinadas questões que dizem respeito ao inatismo (como o problema lógico da aquisição da linguagem), fazendo uso de uma linha argumentativa alicerçada em alguns conceitos-chave selecionados (de fases distintas) da Gramática Gerativo-Transformacional (GGT). Neste sentido, a faculdade da linguagem é apresentada como um sistema inato de computação gramatical – mais especificamente, como um bioprograma de processamento linguístico geneticamente determinado e parcialmente pré-formatado (a Gramática Universal) – e a capacidade cognitiva que os bebês humanos possuem de processar linguisticamente o espectro de dados auditivos (*input* sonoro/acústico) e visuais (*input* luminoso/óptico) que formam a sua “atmosfera” linguística é apresentada como uma dotação biológica que os habilita a internalizar uma língua natural através do mapeamento dos dados linguísticos ambientais. A partir dessas ideias centrais, algumas relações dialéticas são utilizadas como fios condutores da discussão – natureza vs. cultura; órgão da linguagem vs. mecanismos gerais da inteligência; hipótese inatista forte vs. hipótese inatista fraca – com o objetivo de sustentar teoricamente a hipótese de um bioprograma gramatical inato.

No Capítulo 4, são apresentados alguns fatores genéticos que podem estar relacionados com a produção da linguagem a partir de suas bases neurofisiológicas. Inicialmente, são discutidas algumas ideias sobre neuroplasticidade e processamento *on-line* (i.e., em tempo real) de informação linguística, apresentando de que forma a detecção de padrões diferenciais de expressão gênica no tecido cerebral durante a realização de tarefas linguísticas pode ser utilizada para circunscrever e caracterizar os genes que controlam e sustentam a arquitetura e o funcionamento das redes de neurônios responsáveis pela realização da linguagem. Em seguida, são apresentados alguns exemplos de erros inatos do metabolismo (uma subclasse de doenças genéticas) que de alguma forma comprometem a competência e/ou o desempenho linguístico dos indivíduos afetados. O objetivo, obviamente, não é fazer um levantamento exaustivo de todos os componentes genéticos que de algum modo parecem ter ligação com a faculdade da linguagem (um conhecimento do qual ainda estamos muito longe), mas sim apresentar, como modelo de análise, os mecanismos bioquímicos básicos de alguns distúrbios metabólicos causados por mutações genéticas que produzem como um de seus traços fenotípicos déficits linguísticos de natureza variada. Apesar de afetarem a linguagem, tais transtornos metabólicos determinam modificações de caráter mais geral na estrutura cerebral (determinando a qualidade das sinapses, por exemplo) e não é simples delimitar até que ponto os genes envolvidos em tais distúrbios podem ou não ser considerados como fatores diretamente associados à cognição linguística, uma vez que condicionam aspectos estruturais e fisiológicos que subjazem a inúmeros centros cognitivos, e não apenas ao(s) sistema(s) de processamento gramatical. Ainda neste capítulo, são discutidas as causas genéticas de alguns distúrbios (ou transtornos) específicos de linguagem (DEL ou TEL) – uma classe de disfunções que comprometem exclusivamente determinados componentes da linguagem e/ou o seu processo de aquisição, sem que outros sistemas cognitivos sejam afetados. Para tanto, são apresentados e discutidos os resultados de estudos de agregação familiar de DEL (que têm o objetivo de analisar a distribuição fenotípica do transtorno naquelas famílias em que vários indivíduos são afetados por um mesmo subtipo do distúrbio) e de concordância entre gêmeos (cuja finalidade é quantificar estatisticamente os casos em que ambos os gêmeos manifestam o transtorno, comparando-se a taxa de concordância observada entre gêmeos geneticamente idênticos [monozigóticos] com a identificada entre gêmeos fraternos [dizigóticos]).

Dada a sua importância para o estudo das relações entre genes e linguagem, discutimos mais detalhadamente o *FOXP2* – o caso prototípico – no Capítulo 5, onde são

abordados seus aspectos bioquímicos (como localização e estrutura), ontogenéticos (função dos produtos que sintetiza), neurológicos (seu papel na arquitetura cerebral), embriológicos (fases do desenvolvimento em que se expressa mais significativamente) e cognitivo-linguísticos (aspectos da linguagem a ele associados). Como veremos com relação aos membros da família KE – que formam o caso paradigmático –, mutações no *FOXP2* parecem produzir um espectro amplo de problemas linguísticos (alguns deles, bem curiosos). No entanto, o fato de o *FOXP2* ser expresso, por exemplo, no coração de bebês em fase embrionária e no cérebro de animais como ratos e pássaros canoros traz consigo problemas complicados de adequação teórica: afinal, como explicar o fenótipo linguístico a partir de um gene corresponsável por múltiplas funções, expresso em múltiplos órgãos, presente em múltiplas espécies?

A partir de perguntas como essa, realizamos no Capítulo 6 um balanço geral das questões que orientaram o trabalho, retomando alguns pontos, problematizando outros e propondo temas para discussões futuras (agora, sob uma perspectiva mais alinhada ao Programa Minimalista).

Por se tratar de um texto introdutório, buscamos, na medida do possível, oferecer as ferramentas teóricas necessárias para o acompanhamento da discussão, sem deixar o rigor científico de lado, mas também sem que nos perdêssemos em detalhes excessivos durante a exposição. Porém, dada a própria natureza dos assuntos abordados, não nos furtamos a utilizar uma notação bioquímica um pouco mais pormenorizada em alguns pontos do trabalho (em especial, nos capítulos 4 e 5), uma vez que a simplificação excessiva de determinados mecanismos biológicos omitiria detalhes importantes das vias metabólicas que explicam o funcionamento básico de alguns genes (possivelmente) relacionados à emergência da linguagem no cérebro humano. Construir explicações adequadas para o problema da relação entre genes e linguagem representa um desafio imenso para a Linguística e para a Genética, dadas a distância epistemológica de seus métodos de pesquisa e a natureza distinta de seus objetos imediatos de estudo (afinal, genes codificam diretamente cadeias de aminoácidos, e não linguagem); no entanto, se, como disse James Watson, uma parte de nosso destino está escrita nos genes – incluindo a nossa capacidade inata de adquirir uma língua –, é inevitável que tentemos explorar esta zona de interseção de labirintos (o genético e o linguístico), buscando nas profundezas do DNA as diretrizes biológicas responsáveis pela realização do fenômeno da linguagem nos cérebros/mentes de nossa espécie.

CAPÍTULO 2 – A SINGULARIDADE BIOLÓGICA DA FACULDADE DA LINGUAGEM

Como as informações biológicas estão escritas no código genético? Aspectos cognitivos e comportamentais podem ser condicionados geneticamente? Como funcionam determinados sistemas de comunicação animal e por que eles diferem biologicamente da singularidade que caracteriza a linguagem humana? Qual é o potencial simbólico das abelhas? Papagaios realizam Merge? Chimpanzés são capazes de construir sentenças?

2.1. O CÓDIGO DA VIDA

A propriedade mais impressionante da matéria é a vida (ATKINS *et al.*, 2013). É realmente fascinante pensar que fragmentos cósmicos de estrelas e corpos celestes imemoriais, retrabalhados ao longo de bilhões de anos pelas leis da natureza, deram origem, em algum momento, à vida em nosso planeta, formando sistemas físico-químicos auto-organizados capazes de se manter estáveis e de se replicar. Muito mais do que simples aglomerados de matéria, estes “organismos zero” eram sistemas informacionais que possuíam a propriedade fundamental de transferir suas “instruções” de construção e de funcionamento aos seus descendentes através de mecanismos de replicação – uma propriedade que, no limite, derivou todas as formas de vida que habitaram (e habitam) o nosso planeta. Sem esta capacidade autorregulatória de cópia e transferência de dados, a informação da vida se perderia. E sem a modificação destes dados ao longo do tempo, a vida não se transformaria e estaria presa às formas iniciais. Mas para que este conjunto de dados pudesse ser transferido e a vida pudesse ter continuidade, os organismos necessitavam de uma base material capaz de alocar estas informações de modo seguro e estável. Hoje, sabemos que a base física na qual as informações da vida se encontram codificadas corresponde ao material genético – um termo amplo que, normalmente, refere-se às biomoléculas de DNA e/ou de RNA (ambos ácidos nucleicos⁶), mas que também pode se referir aos genes, que são segmentos cromossômicos

⁶ Ácidos nucleicos (como o DNA e o RNA) são macromoléculas formadas por longas cadeias de nucleotídeos ligados entre si (cf. Fig. 2.1).

responsáveis pela síntese de proteínas/cadeias polipeptídicas⁷. Mais adiante veremos com detalhes os mecanismos pelos quais os genes se expressam. No momento, o que nos interessa é perceber como a vida, através de mecanismos de transmissão do material genético, conseguiu atravessar o tempo, transmitindo ao longo das gerações as características de cada espécie e derivando novas formas de organismos. De fato, quando olhamos para o mundo vivo, talvez o que mais nos chame a atenção seja a sua imensa biodiversidade – que não se restringe apenas à multiplicidade de seres macroscópicos que formam os biomas visíveis (como aqueles que ocupam as barreiras de corais ou as florestas tropicais), mas que também diz respeito à profusão de seres que se escondem nos reinos microscópicos da natureza (existem, por exemplo, aproximadamente 700 espécies de bactérias convivendo (em relativa harmonia) na cavidade oral dos seres humanos, totalizando bilhões de seres que dividem este precioso espaço repleto de calor, alimento e água com outras tantas variedades de fungos, vírus e protozoários [DEO & DESHMUKH, 2019])⁸. Entretanto, tão impressionante quanto a riqueza com a qual a vida se manifesta é a unidade que subjaz a todas as suas formas: das

⁷ Em uma definição menos restrita, o termo *gene* se refere a qualquer fragmento de ácido nucleico responsável pela síntese de um produto biologicamente funcional (SNYDER & GERSTEIN, 2003).

⁸ O bioma que vive sobre (e dentro de) nós, seres humanos, é composto por bilhões de seres vivos. “Antony van Leeuwenhoek, o microscopista do século XVII, foi provavelmente o primeiro a ver células bacterianas e reconhecer seu pequeno tamanho, ao dizer: ‘Existe mais vida na boca de um homem que homens em todo o reino’” (GRIFFITHS *et al.*, 2013: 149). Programas como o *The Human Microbiome Project*, um consórcio de pesquisa americano iniciado em 2007 e finalizado em 2016, procuram caracterizar e compreender de que forma este microbioma humano se relaciona com os mecanismos fisiológicos (e mesmo evolutivos) de nossa espécie. “Estima-se que os micro-organismos que vivem dentro e sobre os seres humanos (conhecidos como microbiota) superem as células somáticas e germinativas humanas por um fator de dez. Juntos, os genomas desses simbioss microbianos (definidos coletivamente como microbioma) fornecem características que os humanos não precisaram adquirir evolutivamente por conta própria. Se os humanos são pensados como um composto de células microbianas e humanas, se a paisagem genética humana é pensada como um agregado de genes do genoma humano e do microbioma, e se as características metabólicas humanas são vistas como uma mistura de características humanas e microbianas, então a imagem que surge é a de um ‘supraorganismo’ humano.” (TURNBAUGH *et al.*, 2007: 804). (Texto original: “The microorganisms that live inside and on humans (known as the microbiota) are estimated to outnumber human somatic and germ cells by a factor of ten. Together, the genomes of these microbial symbionts (collectively defined as the microbiome) provide traits that humans did not need to evolve on their own. If humans are thought of as a composite of microbial and humans cells, the human genetic landscape as an aggregate of the genes in the human genome and the microbiome, and human metabolic features as a blend of human and microbial traits, then the picture that emerges is one of a human ‘supraorganism’”).

escamas dos peixes às pétalas das tulipas, do coração das mariposas à íris humana, dos neurônios de primatas ancestrais à raiz dos dentes de dinossauros – todas as informações hereditárias que orientam a construção e o funcionamento da vida estão cifradas em uma linguagem bioquímica que é essencialmente a mesma para todos os seres vivos: o código genético universal.

O código da vida é composto por um alfabeto de apenas quatro componentes bioquímicos: adenina (A), timina (T), guanina (G) e citosina (C). São as bases nitrogenadas que formam os desoxirribonucleotídeos do DNA. Há ainda um quinto componente, a uracila (U), que está presente no RNA.

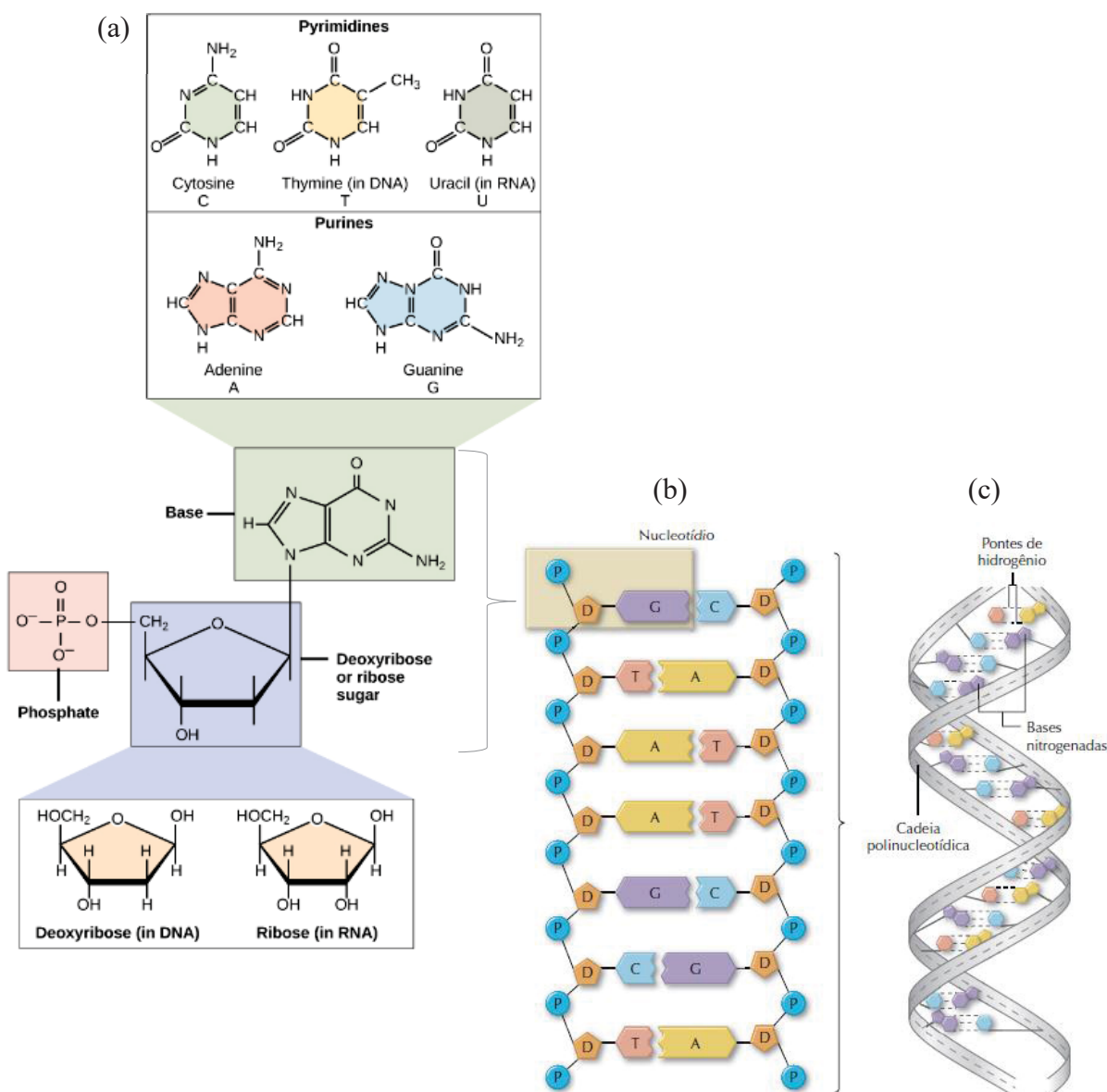


FIGURA 2.1. (a) Fórmula estrutural dos nucleotídeos. Cada nucleotídeo é formado por um grupo fosfato, um açúcar (desoxirribose no DNA; ribose no RNA) e uma base nitrogenada. (b) Representação plana da molécula de

DNA, mostrando as duas cadeias polinucleotídicas unidas por meio de suas bases nitrogenadas: (A)denina se emparelha com (T)imina e (G)uanina se emparelha com (C)itosina. Em destaque, um (desoxirribo)nucleotídeo. (c) Representação tridimensional da dupla-hélice, mostrando as pontes de hidrogênio que mantêm as duas cadeias unidas. Fonte da figura (a): <https://openstax.org/books/biology/pages/3-5-nucleic-acids>. Acesso em 18/12/2020. Fonte das figuras (b) e (c) e dos textos explicativos: AMABIS & MARTHO, 2009: 110.

Como é possível que todas as informações hereditárias estejam codificadas em um sistema constituído por apenas quatro componentes? Para compreendermos o mecanismo básico pelo qual a vida se expressa, vamos fazer uma analogia com o código escrito da língua portuguesa e imaginar que cada letra do nosso alfabeto é codificada por uma sequência específica de bases nitrogenadas (adenina, timina, citosina e guanina). Se cada letra derivasse diretamente de uma única base, teríamos um alfabeto de apenas quatro letras, uma vez que as sequências de bases seriam formadas por somente 1 termo (por exemplo: Adenina $\rightarrow$ letra *a*; Timina $\rightarrow$ letra *b*; Citosina $\rightarrow$ letra *c*; Guanina $\rightarrow$ letra *d*). Já com duas bases, poderíamos construir um alfabeto de 16 letras, pois teríamos 16 sequências distintas:

(1)	Primeira base nitrogenada	Segunda base nitrogenada
	4 possibilidades de escolha: A, T, C, G	4 possibilidades de escolha: A, T, C, G

Princípio Fundamental da Contagem (PFC): $4 \cdot 4 = 16$ sequências de bases (AA; AC; GT etc.).

Pelo mesmo princípio de contagem, caso decidíssemos construir sequências com três bases, teríamos $4 \cdot 4 \cdot 4 = 64$ sequências distintas e, logo, 64 possibilidades de se associar a cada uma delas um símbolo de nosso alfabeto (por exemplo: AAA $\rightarrow$ letra *a*; AAT $\rightarrow$ letra *b*; AAG $\rightarrow$ letra *c*; e assim por diante). Com este sistema, teríamos condições de sobra para representar não apenas as 26 letras de nosso alfabeto (de *a* a *z*), mas também alguns sinais diacríticos (como o *til* e o acento agudo, responsáveis pela marcação de traços fonêmicos) e alguns sinais de pontuação. Desta forma, com apenas quatro unidades básicas tomadas três a três, construímos um sistema de 64 sequências que é capaz de representar todo o alfabeto e, por consequência, todas as sílabas e todas as palavras – no sentido mais usual do termo – do português.

Na natureza, ao invés de codificarem letras que formam palavras, as 64 trincas de bases nitrogenadas codificam aminoácidos, que são as unidades básicas das proteínas (substâncias que, além de determinarem a estrutura das células, estão envolvidas no controle

de todos os processos metabólicos celulares). A Figura (2.2) contém as sequências dos 20 aminoácidos encontrados na natureza.

		Segunda letra					
		U	C	A	G		
Primeira letra	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Fim UAG Fim	UGU } Cys UGC } UGA Fim UGG Trp	Terceira letra	U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }		U C A G
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }		U C A G
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }		U C A G

Abreviaturas dos aminoácidos presentes em proteínas

Ala = alanina	Phe = fenilalanina	Ile = isoleucina	Ser = serina
Arg = arginina	Gly = glicina	Leu = leucina	Thr = treonina
Asn = aspargina	Gln = glutamina	Lys = lisina	Trp = triptofano
Asp = aspartato	Glu = glutamato	Met = metionina	Tyr = tirosina
Cys = cisteína	His = histidina	Pro = prolina	Val = valina

FIGURA 2.2. O código genético. Cada códon (i.e., cada sequência de três bases nitrogenadas do RNA mensageiro [RNAm] no qual são copiadas as informações do DNA) codifica um aminoácido, que será traduzido durante a montagem da proteína. Um segmento de RNAm formado pelos códons AUG-CAC-AGC-GAA, por exemplo, codifica uma sequência de aminoácidos do tipo *Met – His – Ser – Glu*). Fonte: UNIVERSIDADE DE SÃO PAULO. Instituto de Biociências da USP (*Biologia Molecular – O código genético*). Disponível em: https://www.google.com/url?sa=t&source=web&rct=j&url=https://edisciplinas.usp.br/pluginfile.php/3005455/mod_resource/content/1/BiologiaMolecular_texto06final.pdf&ved=2ahUKewiW3uDJg4D4AhWED7kGHTHMCjQQFnoECAkQAQ&usg=AOvVaw1Tuoh0idFrzGdCbdIkdLjB

É com este código e através da cadeia [*um gene* → *um RNA mensageiro* → *uma proteína* → *uma característica*] que os genes controlam a arquitetura e o metabolismo celular

do organismo. Alguns pontos precisam ser esclarecidos com relação ao código da Figura (2.2). A informação codificada no DNA de um gene precisa ser transcrita em uma molécula de RNA mensageiro (RNAm) para que possa ser traduzida pelo sistema bioquímico da célula, transformando-se em proteína (Fig. 2.3). Por isso, o código genético é baseado em tríades nucleotídicas de RNA (também chamadas de códons ou tripletes), e não de DNA (o que pode ser detectado pela presença da uracila (U) na Figura (2.2)). Além disso, todos os aminoácidos, com exceção do triptofano (*Trp*) e da metionina (*Met*), são codificados por mais de uma sequência de bases. No léxico bioquímico da vida, a valina (*Val*), por exemplo, é codificada por quatro “sinônimos”: GUU, GUC, GUA, GUG (neste caso, a sequência inicial GU funciona como a “raiz” do código que representa o aminoácido). A existência de sinônimos faz com que eventuais mudanças na sequência de bases do DNA (mutações) não resultem necessariamente na substituição de um aminoácido por outro e na consequente modificação da proteína. Por isso, diz-se que o código genético é degenerado ou redundante (i.e. existem diferentes códons codificando um mesmo aminoácido).

Outro ponto importante a ser destacado na Figura (2.2) diz respeito às sequências de ignição (AUG) e interrupção (Fim) da síntese proteica. A construção da estrutura primária de uma proteína, ou seja, o enfileiramento de seus aminoácidos, começa quando o sistema de tradução da célula encontra no RNAm a sequência inicializadora AUG e emparelha a metionina, sinalizando o início da polimerização. À medida que os ribossomos – que são as organelas celulares responsáveis pela tradução – vão percorrendo a cadeia de RNAm, ocorre o emparelhamento dos aminoácidos com seus respectivos códons, até o momento em que o sistema de leitura da célula encontra uma das sequências de término (UAA, UAG ou UGA) e a síntese é interrompida, liberando a proteína/polipeptídeo. Sem estes pontos finais geneticamente programados, os produtos proteicos seriam disfuncionais e a vida, da forma como a conhecemos, provavelmente seria inviável.

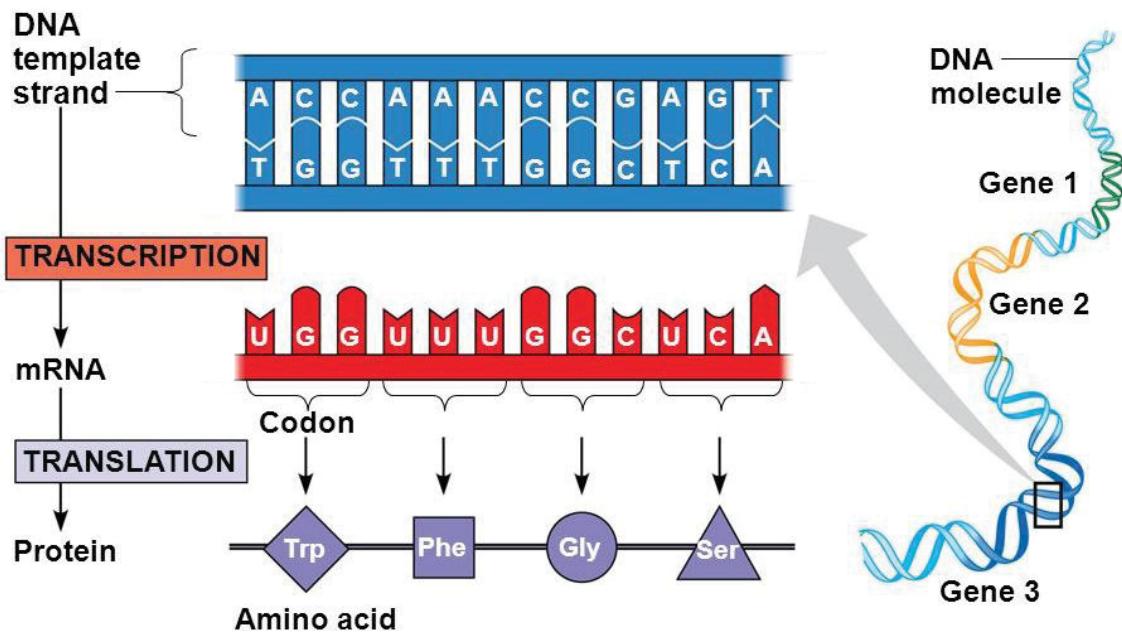


FIGURA 2.3. O dogma central da biologia molecular. “Para cada gene, uma das fitas de DNA funciona como um molde (*DNA template strand*) para a transcrição de moléculas de RNA, como o RNAm. A regra de pareamento de bases nitrogenadas para a síntese de DNA [adenina (A) ↔ timina (T); citosina (C) ↔ guanina (G)] também guia a transcrição, mas a uracila (U) substitui a timina (T) durante a síntese do RNA. Durante a tradução, o RNAm é lido como uma sequência de trincas [ou tripletes] de nucleotídeos, chamadas de códon. Cada códon especifica um aminoácido que será adicionado à cadeia polipeptídica em crescimento.”⁹ Figura e texto extraídos e adaptados de CAMPBELL *et al.*, 2020: 340.

A existência de um código genético universal que codifica os mesmos componentes bioquímicos e que é lido da mesma forma pela engenharia celular de todos os seres vivos¹⁰ é uma demonstração poderosa da unidade da vida. Se ampliarmos a analogia que fizemos entre o código genético e o código da escrita, podemos considerar que cada espécie de nosso planeta – na realidade, cada ser vivo que existiu e que existe – é um texto biológico que deriva da mesma linguagem bioquímica universal. Algumas receitas moleculares, aliás, são amplamente compartilhadas, como é o caso do citocromo c, um pigmento celular respiratório envolvido no transporte de elétrons. O bioprograma de construção desta proteína está presente

⁹ “For each gene, one DNA strand functions as a template for transcription of RNAs, such as mRNA. The base-pairing rules for DNA synthesis also guide transcription, except that RNA is made with uracil (U) instead of thymine (T). During translation, the mRNA is read as a sequence of nucleotide triplets, called codons. Each codon specifies an amino acid to be added to the growing polypeptide chain.”

¹⁰ Algumas sequências de bases são lidas de maneira diferente nas mitocôndrias e nos cloroplastos, constituindo exceções à universalidade do código genético.

no material genético de seres que vão desde os fungos do gênero *Candida* até os *Homo sapiens*:

A homologia do DNA é geralmente surpreendente; por exemplo, o DNA codificador e a sequência de aminoácidos da proteína citocromo c de transporte de elétrons são homólogos em inúmeros organismos no planeta, abrangendo bactérias, fungos, vermes, insetos, mamíferos e assim por diante. Esse tipo de revelação, em conjunto com a demonstração da grande homologia dos processos bioquímicos que atuam nas células, tem levado a uma nova consciência de que os seres humanos na verdade são "primos" de todas as formas de vida na Terra (GRIFFITHS *et al.*, 2013: 14).

A sequência de nucleotídeos do gene que codifica o citocromo c em diferentes espécies, entretanto, não é exatamente a mesma. Existem pequenas diferenças, fruto de mutações que foram se acumulando ao longo da evolução. Esta taxa de variação, hoje, é utilizada no cálculo de relações evolutivas e na construção de árvores filogenéticas: quanto maior for a homologia das sequências nucleotídicas deste gene, maior será o grau de parentesco entre duas espécies (mais precisamente, entre duas categorias taxonômicas, já que podemos pensar em termos de gêneros, famílias, ordens, classes etc.). A homologia do gene codificador desta proteína é maior entre seres humanos e cavalos do que entre seres humanos e atuns, por exemplo (GRIFFITHS *et al.*, 2013: 14). Porém, apesar de variar ao longo do tempo, o programa de construção do citocromo c continua sendo escrito com o mesmo código bioquímico universal.

Mas, afinal, o que a natureza pode escrever com este código? Sabemos há algum tempo que traços fenotípicos¹¹ como a textura das sementes de ervilha (*Pisum sativum*), a cor dos olhos da mosca-da-fruta (*Drosophila melanogaster*) ou o número de dedos dos seres humanos (*Homo sapiens*) são características condicionadas geneticamente. Sabemos também que os genes são diretamente responsáveis pela arquitetura da vida (genes *Hox*, por exemplo, são corresponsáveis pelo desenvolvimento padrão do eixo corporal dos animais, definindo,

¹¹ O fenótipo primário de um gene é a proteína/polipeptídeo que ele produz. No entanto, quando nos referimos a um traço fenotípico, normalmente estamos fazendo referência a uma característica visível ou detectável do organismo e precisamos levar em consideração tanto a constituição genética do indivíduo (genótipo) quanto as possíveis influências das condições ambientais. Assim, Fenótipo = Genótipo + Fatores ambientais. A altura dos seres humanos e a cor das flores de hortênsia, por exemplo, são traços fenotípicos determinados geneticamente que podem ser influenciados por fatores ambientais como a dieta nutricional ou a acidez do solo, respectivamente.

dentre outras coisas, o local correto em que os membros – pernas, antenas, asas – devem se desenvolver [McGINNIS *et al.*, 1984]) e pelo seu funcionamento adequado (mutações do gene que codifica a enzima fenilalanina hidroxilase (PAH), por exemplo, impedem que o organismo converta a fenilalanina contida nos alimentos no aminoácido tirosina, causando a fenilcetonúria (PKU) em humanos (cf. Fig. 3.3 e a seção 4.2.2), uma doença grave que exige uma dieta especial e que pode ser detectada pelo teste do pezinho em bebês). Mas e se formos além? E se pensarmos na genética dos comportamentos e dos instintos?

Aranhas tecem teias com padrões geométricos típicos de cada espécie e apresentam padrões motores extremamente intrincados de colocação e amarração dos fios (ADES, 1995). Em um determinado momento de seu ciclo de vida, salmões atravessam os mares para se reproduzir exatamente no rio em que nasceram – e grande parte não sobrevive a este impulso migratório de vida e morte (COOKE *et al.*, 2011; THORPE, 1988). Com aproximadamente 5-6 semanas de vida, gatos domésticos já fazem covas para enterrar seus dejetos e começam a ensaiar os primeiros movimentos de caça (BATESON, 2014). Quais seriam as origens do projeto de tecelagem das aranhas, do sistema de localização espacial dos salmões e dos impulsos predatórios dos felinos¹²? O que, no comportamento destes animais, seria determinado e controlado pelo programa genético da espécie e o que seria fruto de alguma forma de aprendizado? A discussão sobre a natureza instintiva ou adquirida de determinados comportamentos não é nova. Ainda no século XIX, Charles Darwin já havia estendido suas ideias sobre evolução à dimensão comportamental dos seres vivos:

Uma primeira, e notável, contribuição de Darwin reside na generalização dos princípios da seleção natural ao comportamento instintivo. O princípio é simples: os traços comportamentais, como os anatômicos e fisiológicos, variam entre indivíduos, transmitem-se por hereditariedade e se tornam mais frequentes na medida em que proporcionem aos indivíduos uma capacidade maior para enfrentar os desafios ambientais e para se reproduzir. Darwin (1859) [*A origem das*

¹² Apesar das semelhanças existentes entre os padrões motores demonstrados pelos gatos durante as brincadeiras sociais na infância e durante as atividades de caça, Bateson (2014: 19) questiona a hipótese de que o ato de brincar seja um estágio necessário de preparação da atividade predatória: “[...] Gatos ‘Kaspar Hauser’ que foram criados em isolamento social, sem oportunidades de experiência visual e sem acesso a comportamentos relacionados a brincadeiras, [não obstante a falta de estímulos], mostraram respostas predatórias ‘normais’ quando apresentados a um manequim móvel semelhante a uma presa [...]” (Texto original: “[...] ‘Kaspar Hauser’ cats which were reared in social isolation and without opportunities for visual experience, let alone play behaviour, nonetheless showed ‘normal’ predatory responses when presented with a prey-like moving dummy at 11 weeks of age [...]”).

espécies] aplica a ideia a instintos impressionantes como a tendência do cuco europeu em colocar seus ovos em ninhos alheios, o comportamento das formigas que usam formigas de outras espécies como escravas para a realização das tarefas do ninho e a perfeição hexagonal dos alvéolos nos favos de abelhas. Em cada caso, variantes individuais poderiam ter sido selecionadas, ao longo das gerações, em função da vantagem reprodutiva, ganhando predominância na população (ADES, 2009: 44).

De fato, apesar do papel fundamental que a aprendizagem, a experiência e a memória exercem no refinamento de determinados comportamentos instintivos – a taxa de eficiência da atividade predatória de animais caçadores (como aranhas e gatos) tende a ser menor entre indivíduos muito jovens, por exemplo –, parece ser arriscado assumir que rotinas de execução tão complexas quanto a construção de uma colmeia ou a odisseia reprodutiva dos salmões possam emergir do cérebro desses animais sem a presença de sistemas inatos de diretrizes mentais que os oriente. Se os conjuntos neuronais de cada espécie são codificados e herdados geneticamente, as funções associadas a eles também o são – o que inclui os comportamentos instintivos que deles derivam.

Assumir que os *softwares* de construção de teias e de colmeias já vêm instalados no cérebro de aranhas e de abelhas tem implicações filosóficas e científicas importantes. Ainda que estes conjuntos de instruções possam não estar totalmente configurados ao nascimento e apresentem espaços para a aprendizagem, estamos assumindo que mesmo cérebros simples como o de uma abelha – que mede 1 mm³ e possui cerca de 1 milhão de neurônios (BRANDT *et al.*, 2005: 2) – possuem sistemas de cognição pré-formatados que permitem a execução de comportamentos instintivos. Com isso, estamos questionando a ideia de que os seres vivos nascem sem nenhum tipo de conhecimento. Aranhas e abelhas não são tábulas rasas: seus comportamentos instintivos fazem parte de sua história evolutiva natural e estão fundados em subsistemas de neurônios geneticamente determinados. Se, durante a gametogênese, ocorrem mutações aleatórias em genes neuromodeladores e se, após a fusão dos gametas na reprodução¹³, estas mutações resultam em sistemas neuronais e em comportamentos instintivos que representam vantagens adaptativas para o indivíduo, é de se esperar que estes genes – e, indiretamente, os neurônios e instintos por eles condicionados – sejam transmitidos

¹³ Estamos nos referindo à fusão de gametas (i.e. à fecundação) como uma fase do mecanismo biológico de transmissão de genes. Há, entretanto, outras modalidades de reprodução e de combinação gênica. Abelhas, por exemplo, apresentam partenogênese: óvulos fecundados por espermatozoides dão origem a fêmeas; óvulos não fecundados dão origem a machos (zangões).

às gerações subsequentes. Como Darwin já havia observado, a seleção natural age também sobre os comportamentos.

Mas o que dizer sobre performances comunicativas? A capacidade de emissão e recepção de mensagens é um mecanismo indispensável à sobrevivência e atravessa todo o mundo vivo – incluindo aí todas as células do nosso corpo, que neste exato momento estão enviando e recebendo bilhões de informações bioquímicas a fim de nos manter vivos. Na próxima seção, analisaremos como alguns animais se comunicam e discutiremos as semelhanças e diferenças entre estes sistemas de comunicação e a linguagem humana.

2.2. COMUNICAÇÃO ANIMAL

2.2.1. Abelhas dançarinas

Além de construírem instintivamente complexos favos hexagonais, abelhas demonstram uma série de comportamentos derivados de sua capacidade comunicativa, como a organização prodigiosa de suas colônias e a divisão sistemática de tarefas. Em um texto, hoje clássico, intitulado “Comunicação animal e linguagem humana”, Émile Benveniste (1952 [1976]), baseado nos estudos entomológicos e etológicos de Karl von Frisch¹⁴, analisa e problematiza o engenhoso mecanismo de comunicação que as abelhas-europeias (*Apis mellifera*) utilizam durante seu forrageamento. O problema enunciado por Benveniste é, no mínimo, instigante: após a localização do alimento (pólen, néctar ou qualquer outra substância açucarada), a abelha exploradora retorna à colmeia e avisa às companheiras sobre o doce achado. Em seguida, abelhas distintas (e sem o acompanhamento da exploradora inicial) se dirigem exatamente à fonte de alimento, indicando que houve efetivamente algum tipo de transmissão de mensagem entre a abelha exploradora inicial e as demais. Mas como?

Realizando experiências com colmeias transparentes, Karl von Frisch constatou que as abelhas se comunicam através de um sistema simbólico formado por “passos de dança” que codificam informações espaciais e possibilitam a transmissão de coordenadas geográficas. O

¹⁴ Karl Ritter von Frisch (1886 – 1982) foi um dos fundadores da Etologia, ciência que faz o estudo comparado do comportamento dos animais, de seus hábitos e de sua acomodação às condições ambientais. Por seus estudos, recebeu o Prêmio Nobel de Fisiologia ou Medicina de 1973.

conjunto de sinais que as abelhas utilizam para simbolizar o espaço possui dois elementos básicos de significação: danças que delineiam círculos no ar e danças que desenhavam bidimensionalmente uma espécie de “oito” (Fig. 2.4a). Neste último movimento, as abelhas complementam a manobra com uma vibração do abdômen (conhecida como *wagging-dance* ou “dança do ventre”). Quando a distância da colmeia até a fonte de alimento se encontra dentro de um raio de aproximadamente 100 metros, a abelha que a encontrou dança em círculos diante de suas companheiras, que se espalham centrifugamente em volta do ninho. Não há nesta dança circular indicações sobre direção e sentido, apenas a transmissão da informação de que existe alimento nos arredores. Quando o alimento se encontra a uma distância de 100 metros a seis quilômetros do ninho, a abelha exploradora desenha sucessivos “oitos” na parede da colmeia, os quais possuem três informações conjugadas: (i) a frequência (f) dos movimentos indica a distância (d) da colmeia à fonte – quanto maior é a distância, menor é o número de voltas que a abelha realiza por período de tempo, ou seja, mais lenta é a dança. No balé das abelhas, d e f são duas grandezas inversamente proporcionais: $d \propto f^{-1}$. Note que, quanto mais longe o alimento se encontra do ninho, mais tempo a abelha leva para encontrá-lo. Este intervalo de tempo mais extenso é “traduzido” pela abelha sob a forma de uma performance mais lenta durante a dança, indicando que estes insetos talvez consigam codificar não apenas dados espaciais, mas também informações temporais através da modulação da velocidade. Uma vez que o sistema cognitivo de mapeamento das abelhas é capaz de processar informações de espaço (s) e de tempo (t) de forma complexa e instintiva, é provável que estes insetos consigam calcular com eficiência as relações que são derivadas destas duas grandezas fundamentais – como velocidade ($v = \Delta s / \Delta t$), frequência ($f = n / \Delta t$) e período ($T = f^{-1} = \Delta t / n$)¹⁵ – durante a (de)codificação de sua mensagem. Assim, quanto menor for a velocidade de realização do “oito” simbólico, por exemplo, menor será a frequência da dança e maior será a distância do ninho ao alimento; (ii) a inclinação do eixo imaginário do “oito” indica a direção e o sentido em que as abelhas receptoras da mensagem devem seguir – um cálculo espacial que é realizado através da triangulação da colmeia, do sol e da fonte de alimento (Fig. 2.4b); (iii) a frequência com que a abelha “rebola” é um indicativo da qualidade do alimento – assim, quanto melhor é a fonte de recursos, mais “agitada” é a vibração de seu ventre, numa curiosa correspondência entre “grau de qualidade” (ou mesmo “grau de satisfação”) e “grau de movimento”. A complexidade das informações codificadas e

¹⁵ v : velocidade; Δs (deslocamento espacial) = $s_f - s_i$; Δt (intervalo de tempo) = $t_f - t_i$; f : frequência; T : período; n : número de ciclos.

decodificadas nos passos de dança das abelhas indica que o comportamento comunicativo destes insetos, muito provavelmente, possui uma base instintiva. Se a *Apis mellifera* é capaz de ler, calcular e (de)codificar simbolicamente as coordenadas do espaço, isso só é possível graças a um processador neurobiológico específico e ao *software* de mapeamento que nele veio instalado.

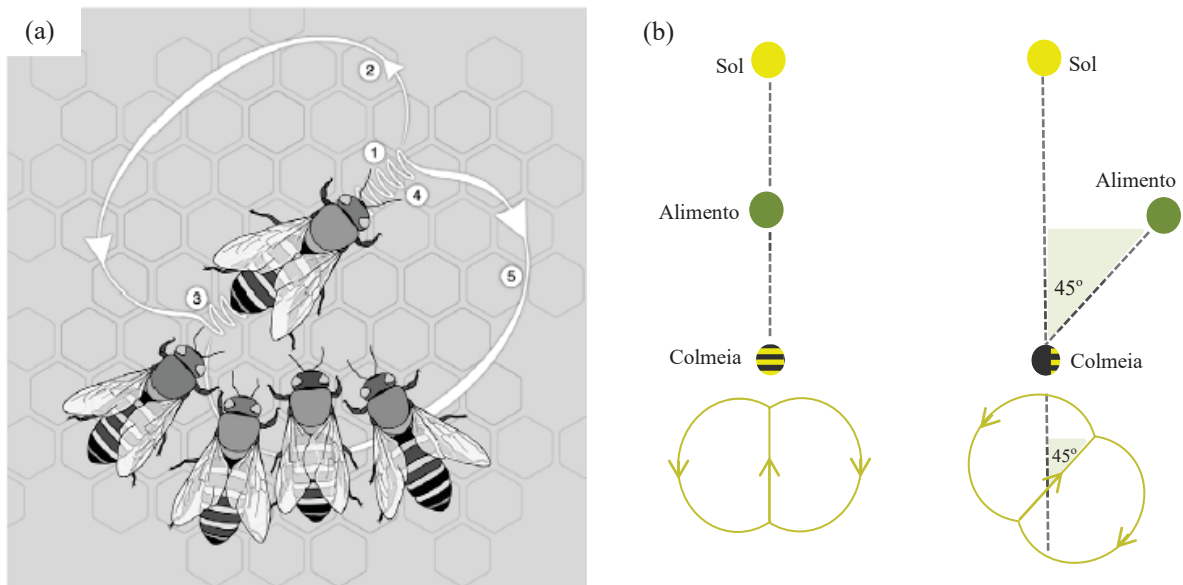


FIGURA 2.4. A dança das abelhas. Em **(a)**, a abelha emissora dança diante das companheiras (receptoras). A dançarina se desloca em linha reta vibrando o abdômen (1) e vira para um dos lados (2), realizando um semicírculo e voltando ao ponto inicial do movimento (3), onde dá início a uma nova corrida em linha reta enquanto vibra o abdômen (4) para então virar para o outro lado (5) e voltar ao começo, onde reinicia o ciclo. Quanto melhor é a fonte de alimento, mais corridas ela executa e maior é o estado de agitação de sua performance. Durante a corrida, a dançarina produz sons, fluxos de ar e vibrações na colmeia, além de liberar substâncias químicas. As seguidoras tocam a dançarina diretamente com suas antenas durante a dança. No entanto, a importância dos diferentes estímulos para a transferência de informações espaciais da dançarina para as seguidoras ainda não é bem compreendida. Além de fornecer informações de localização da fonte de alimento, esses componentes podem ter uma importante função psicológica para as abelhas receptoras, uma vez que a sinalização multicomponente pode melhorar a detecção do sinal e melhorar o aprendizado das informações (texto extraído e adaptado de GRÜTER & FARINA, 2009: 243). **(b)** O ângulo de inclinação do eixo central da dança em relação aos favos da colmeia codifica o ângulo formado pela triangulação entre a colmeia, a fonte de alimento e o Sol. Fonte da figura (a): GRÜTER & FARINA, 2009. Figura (b) adaptada de <https://www.newscientist.com/article/mg20327262-400-rethinking-the-bees-waggle-dance/amp/>. Acesso em 13/01/2021.

Não há dúvidas quanto ao grau de sofisticação do sistema de comunicação da *Apis mellifera*. No entanto, “pode ser útil assinalar de leve aquilo em que ela [a “linguagem” das

abelhas] *é ou não uma linguagem, e o modo como essas observações sobre as abelhas ajudam a definir, por semelhança ou por contraste, a linguagem humana*” (BENVENISTE, 1976 [1952]: 64). Quanto às semelhanças, Benveniste assim as elenca:

As abelhas se mostram capazes de produzir e compreender uma verdadeira mensagem, que encerra inúmeros dados. Podem, pois, registrar relações de posição e distância; podem conservá-las na ‘memória’; podem comunicá-las simbolizando-as por diversos comportamentos somáticos. O fato notável consiste inicialmente em que manifestam aptidão para simbolizar: há, mesmo, uma correspondência ‘convencional’ entre seu comportamento e o dado que traduz. Essa correspondência é percebida pelas outras abelhas nos termos em que lhes é transmitida e se torna em motor de ação. Até aqui encontramos, nas abelhas, as próprias condições sem as quais nenhuma linguagem é possível – a capacidade de formular e interpretar um ‘signo’ que remete a uma certa ‘realidade’, a memória da experiência e a aptidão para decompô-la¹⁶ (BENVENISTE, 1976: 64).

O mecanismo simbólico das abelhas, portanto, codifica os dados objetivos da realidade em “comportamentos somáticos” convencionalizados. Estes signos, no entanto, precisam ter significação constante, sob o risco de a mensagem não ser compreendida pelos outros membros da colônia. Assim, o sistema de comunicação das abelhas é fundado na fixidez dos signos, isto é, na significação constante de suas danças, que sempre se referem ao mesmo “assunto” (a localização do alimento) e são construídas com o mesmo conjunto estrito de movimentos. No mundo das abelhas, não há espaço para danças polissêmicas: um movimento em “oito” em determinada frequência sempre terá o mesmo significado. Também não há espaço para a criação de novos movimentos simbólicos: voos não convencionais (como ziguezagues) não codificam mensagens. Além disso, as mensagens transmitidas pelas abelhas são indecomponíveis: movimentos inacabados não possuem significação.

¹⁶ A aptidão para a comunicação deriva de uma série de competências biocognitivas codificadas no genoma das abelhas: “Por causa da sua importância para a agricultura e a ciência, a abelha foi selecionada para sequenciamento pelo *National Genome Research Institute* dos Estados Unidos. A sequência completa do genoma foi publicada em 2006 e forneceu numerosas informações sobre o comportamento, a função neurológica, a ecologia e a evolução das abelhas. Seu genoma é composto por 10.157 genes que incluem cerca de 236 milhões de pares de bases de DNA. Com base na informação genômica, foram identificados 36 genes que codificam neuropeptídeos, proteínas do cérebro, que afetam o comportamento e a memória. Mais de 3.000 genes estão ativos no cérebro de uma abelha, e os pesquisadores localizaram regiões reguladoras que controlam a expressão de muitos destes genes, incluindo alguns que são importantes no desenvolvimento do comportamento de procura por alimento. Comparada com outros insetos, as abelhas têm muito mais genes codificadores de receptores olfatórios, explicando seu sofisticado sistema químico de comunicação.” (PIERCE, 2016: 853).

Já a linguagem humana, por contraste, possui uma base fisiológica vocal (mas também gestual, como nas diversas línguas de sinais), possibilita o diálogo livre (que é inexistente entre as abelhas, onde não há troca de turnos), permite a plurissignificação e a construção de sentidos figurados (enquanto as abelhas só codificam e entendem o sentido “literal” dos signos) e oferece condições para que a comunicação faça referência a um dado exclusivamente linguístico (no caso das abelhas, a transmissão de mensagens só é possível quando a abelha vê o alimento, sendo impossível a retransmissão direta da mensagem veiculada pela abelha exploradora – ou seja, falta ao sistema comunicativo das abelhas aquilo que é um dos pontos centrais da definição de linguagem: a substituição simbólica plena da experiência que, uma vez codificada, pode ser retransmitida indefinidamente no tempo e no espaço. “*A abelha não constrói uma mensagem a partir de outra mensagem*” [Ibid.: 65]). Além disso, existe um potencial infinito de conteúdos e assuntos que podem ser processados e veiculados pelo sistema linguístico humano (ao passo que as abelhas só realizam monólogos sobre um único tema e um único dado da realidade: a direção e o sentido do alimento encontrado) e um potencial infinito de formas linguísticas (sentenças) que podem ser construídas através da combinação de elementos finitos segundo regras específicas também finitas (o que marca a infinitude discreta de nossas línguas). Os humanos são linguisticamente criativos; as abelhas, não. Por fim, as línguas humanas são duplamente articuladas: suas formas linguísticas podem ser decompostas em elementos significativos menores (morfemas) que, por sua vez, podem ser reduzidos a constituintes articulatórios sem significação (fonemas). Já o sistema de comunicação das abelhas não se deixa reduzir a “*elementos identificáveis e distintivos*” (Ibid.: 66).

Assim, dado este conjunto de diferenças, Benveniste conclui que o sistema comunicativo das abelhas se configura não como uma (proto)linguagem, mas sim como um código de sinais e ressalta, pelo contraste, aquilo que o diferencia diametralmente da linguagem humana: “*a fixidez do conteúdo, a invariabilidade da mensagem, a referência a uma única situação, a natureza indecomponível do enunciado, a sua transmissão unilateral*” (Ibid: 67).

Apesar dos contrastes entre o código das abelhas e a linguagem humana, a existência de um mecanismo mental que é capaz de processar os dados ambientais – como a posição relativa do Sol, por exemplo – e de transformá-los em signos aponta para uma característica fundamental do aparato cognitivo das abelhas – e que é parte fundamental do sistema linguístico humano: o pensamento simbólico.

Desde os anos 1920, experimentos realizados com abelhas da espécie *Apis mellifera* têm demonstrado a habilidade desses insetos em realizar discriminações simples. A tarefa mais comum idealizada pelos pesquisadores consiste em sinalizar com cores diferentes as entradas de dois locais que possuem bebedouros abastecidos com líquidos distintos – um com xarope açucarado e outro apenas com água. Após algumas horas de “tentativa e erro”, as abelhas aprendem que só uma das cores está associada ao xarope e sistematicamente se dirigem ao bebedouro correto, mesmo quando a posição dos bebedouros é invertida (tal procedimento de inversão visa eliminar a hipótese de que as abelhas conseguem realizar a discriminação adequada fazendo uso apenas da memória espacial). A *Apis mellifera*, inclusive, mostrou-se capaz de realizar discriminações entre padrões visuais mais complexos do que cores e figuras geométricas simples. Em estudo de 2013, abelhas desta espécie demonstraram que conseguem diferenciar quadros de Picasso e de Monet durante a realização do mesmo tipo de experimento.

Assim, nosso estudo contribui para o crescente corpo de evidências de que insetos como as abelhas têm a capacidade de aprender, reter, classificar e processar informações visuais de uma maneira que não é prevista por simples respostas mecanicistas ou elementares a estímulos [...]. É claro que o fato de as abelhas poderem discriminar pinturas não implica que elas realmente reconheçam um estilo artístico *per se* ou interpretem o conteúdo da obra de arte em um nível semântico de maneira semelhante aos humanos. Mas sugere que a discriminação de estilos artísticos não é uma função cognitiva superior exclusiva de seres humanos, mas pode ser simplesmente uma extensão da capacidade que os animais possuem (de insetos a humanos) de extrair e categorizar as características visuais de imagens complexas. (WU *et al.*, 2013: 54)¹⁷.

Além de constatarem o elevado grau de acuidade visual da *Apis mellifera*, estes experimentos de *discriminação simples* evidenciam a capacidade que esses insetos possuem de processar signos do tipo [*cor/figura A* ↔ *xarope*], [*cor/figura B* ↔ *água*] e apontam para a existência de um pensamento (pré-)simbólico de base, sem o qual estas relações não poderiam ser construídas/aprendidas. Experimentos de *discriminação condicional* reforçam ainda mais

¹⁷ “Thus, our study contributes to the growing body of evidence that insects like honeybees have the ability to learn, retain, classify and process visual information in a way that is not predicted by simple mechanistic or elemental responses to stimuli [...]. Of course, the fact that bees can discriminate paintings does not imply that bees actually recognize an artistic style *per se* or interpret art content at a semantic level in a similar way to humans. But it suggests that discrimination of artistic styles is not a higher cognitive function that is unique to humans, but may simply be due to the capacity of animals, from insects to humans, to extract and categorize the visual characteristics of complex images.”

a proposta de que as abelhas possuem uma capacidade elevada de compreender relações arbitrárias entre dois elementos distintos. Em um destes experimentos, Moreno *et al.* (2012) testaram a capacidade de discriminação condicional de abelhas das espécies *Apis mellifera* e *Melipona rufiventris* (um tipo de abelha sem ferrão, conhecida popularmente como tujuba), fazendo-as passar por uma espécie de mini-labirinto formado por três câmaras conectadas em forma de Y e sinalizadas por padrões geométricos e cores (Fig. 2.5). Quando a câmara de entrada estava sinalizada com listras verticais, o xarope se encontrava na câmara marcada com a cor amarela; quando as listras que sinalizavam a câmara de entrada eram horizontais, o xarope estava na câmara marcada pela cor azul. A ideia é simples: a escolha da cor correta (aquela que leva ao alimento) é determinada/condicionada pelo padrão de listras antecedente. Trata-se da construção de uma relação totalmente arbitrária entre padrões geométricos e cores: [listras verticais – cor amarela], [listras horizontais – cor azul] – tão arbitrária quanto a relação entre a palavra *mel* e o objeto (ou conteúdo referencial) que ela aponta no mundo. Após seis dias de treinos, as abelhas da espécie *Apis mellifera* conseguiram acertar aproximadamente 70% das discriminações condicionais, mesmo quando os bebedouros não estavam mais presentes, o que refuta a hipótese de que as abelhas pudessem estar se orientando pelo olfato. Entre as tujubas, o índice de acerto foi de aproximadamente 50%, o mesmo que ocorreria ao acaso, indicando que esta espécie, aparentemente, não é capaz de aprender/construir relações arbitrárias.

Apesar de demonstrarem uma aptidão para o pensamento simbólico, as abelhas-europeias, segundo Moreno *et al.* (2012), não são capazes de construir novas relações arbitrárias a partir de relações aprendidas. Uma abelha que tenha aprendido a escolher a cor amarela após atravessar uma câmara sinalizada por listras verticais, por exemplo, não consegue aplicar a mesma associação entre padrão geométrico e cor (*listras verticais* ↔ *cor amarela*) quando precisa realizar a operação inversa para encontrar o alimento, ou seja, quando precisa atravessar um labirinto em que o alimento se encontra em uma câmara de listras verticais *antecedida* por uma câmara marcada pela cor amarela. Tudo indica que a discriminação condicional é memorizada pela abelha na ordem em que foi apresentada (*listras verticais* → *cor amarela*) e não pode ser manipulada: ao atravessar uma sala sinalizada pela cor amarela, a abelha escolhe o padrão geométrico da sala seguinte ao acaso, mostrando que ela não é capaz de realizar inferências transitivas que lhe possibilitem “inverter” a associação aprendida (i.e., ela é incapaz de inferir e construir a relação simétrica *cor amarela* → *listras verticais*).

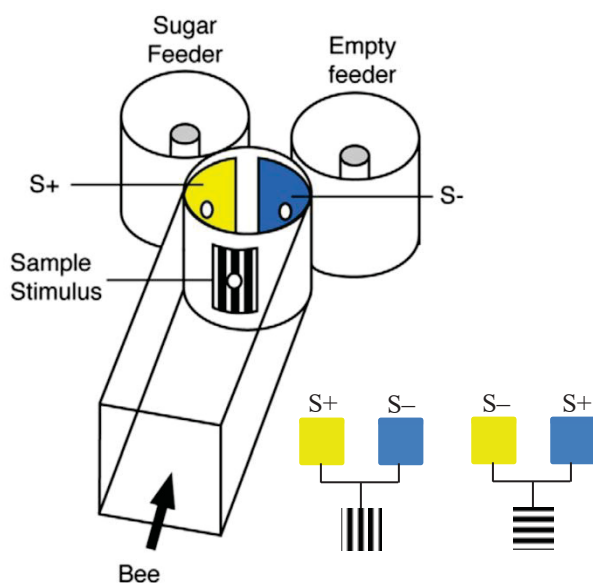


FIGURA 2.5. Esquema de um labirinto-Y usado para investigar a aprendizagem de abelhas com relação à capacidade de realizar discriminações condicionadas e construir relações simbólicas arbitrárias. “A abelha entra em um túnel, voa através da entrada do labirinto, marcada com o estímulo condicionante (*sample stimulus*), e então atravessa uma de duas entradas, cada uma delas marcada com uma cor distinta (*comparison stimulus*: estímulo condicionado). Uma escolha correta (S+) leva a abelha a um bebedouro contendo solução açucarada. Uma escolha incorreta a leva a um bebedouro vazio ou contendo apenas água (S-)”. Imagem e texto extraídos de MORENO *et al.*, 2012: 2.

Uma vez que os experimentos mostram que, após a aprendizagem, o comportamento das abelhas é tal que um padrão geométrico X as “conduz” à cor Y, mas a cor Y não as “conduz” necessariamente ao padrão geométrico X, parece um pouco precipitado considerar que o par *padrão geométrico X – cor Y* se constrói como um verdadeiro signo na mente das abelhas. Um signo é uma realidade de mão-dupla: aquilo que representa (*significante*) leva àquilo que é representado (*significado*) e vice-versa. No caso das abelhas-europeias, os pares de sinais que elas abstraem dos experimentos são organizados em um único sentido e não podem ser manipulados inversamente (simetricamente). Dado que a relação arbitrária entre os sinais é aprendida e memorizada pela *Apis mellifera* levando-se em conta a ordem em que aparecem, talvez falte a esta espécie o passo seguinte rumo à simbolização plena, que é a transformação de uma *sequência* (ou arranjo) de *dois sinais relacionados* em um *único sinal*, mas *de face dupla* – i.e., em um *signo*.

2.2.2. Papagaios, beija-flores e pássaros canoros

Embora possam ser também a manifestação instintiva de um bem-estar psicológico – como ocorre aos galos quando veem o Sol despontando no horizonte ou às gaivotas quando sentem as correntes de ar do crepúsculo –, os cantos das aves têm uma função primordialmente comunicativa (que pode ser simplesmente avisar sobre o dia que nasce ou cantar sobre o dia que morre). Três grupos especiais de aves – pássaros canoros (como o

mandarim-diamante, *Taeniopygia guttata*), psitacídeos (como o papagaio-cinzento, *Psittacus erithacus*) e beija-flores – utilizam combinações complexas de sons para realizar diferentes funções ecológicas: demarcar território, avisar sobre a presença de alimento ou de predadores e atrair parceiros sexuais. Psitacídeos são exímios imitadores e podem dominar um inventário de centenas de (combinações de) sons; pássaros canoros e beija-flores, por sua vez, normalmente se dedicam a uma única melodia, com pequenas variações temáticas.

Diferentemente de outros grupos de aves, que apresentam apenas as vocalizações inatas típicas de cada espécie, os membros destas três ordens – que, em conjunto, formam um grupo que aqui chamaremos simplesmente de *aves canoras* – são considerados aprendizes vocais, pois são capazes de imitar sons e aprender melodias. Três aparelhos evolutivos permitiram que estas três ordens de aves alcançassem um sistema extremamente refinado de manipulação de sons (JARVIS, 2004: 752-755): (i) um aparelho fonador (siringe) finamente controlado, com o qual estas aves instanciam e modulam mecanicamente as combinações de sons; (ii) um conjunto de núcleos cerebrais especialmente dedicados à (de)codificação, processamento e memorização de melodias, harmonias e padrões sonoros (*auditory regions*) e (iii) um complexo de núcleos cerebrais responsáveis pela programação da articulação dos sons que são vocalizados (*vocal regions*). Para neurobiólogos como Erich Jarvis (2004), são núcleos cerebrais como os de (iii), extremamente raros na natureza, que tornam alguns grupos de seres vivos – entre eles, as aves canoras e os seres humanos – aptos a desenvolverem a chamada “aprendizagem vocal”: “*Aprendizagem vocal é a habilidade de adquirir vocalizações por imitação ao invés do instinto*” (JARVIS, 2004: 749)¹⁸. Mesmo que ao longo de suas linhas evolutivas aves canoras e seres humanos tenham adquirido de forma independente a capacidade de vocalizar, o mapeamento neurológico dos núcleos de processamento de sons destes três grupos de aves e a comparação com o(s) núcleo(s) de produção da fala dos seres humanos têm revelado algumas homologias importantes entre as regiões cerebrais responsáveis pela programação e articulação dos sons (vocalizações) destas espécies. Muitas vezes, este mapeamento é feito comparando-se cérebros lesionados de aves e de seres humanos que apresentam algum déficit motor facial:

[...] lesões nas estruturas HVC e RA [dois dos núcleos cerebrais vocais dos pássaros canoros], particularmente aquelas localizadas no hemisfério esquerdo de canários, causam déficits semelhantes àqueles encontrados após lesões no córtex motor da face humana e que resultam em mudez para vocalizações aprendidas, i.e., para a fala [e para os cantos, no caso dos

¹⁸ “Vocal learning is the ability to acquire vocalizations through imitation rather than instinct.”

canários]. Sons inatos, como choros e gritos, entretanto, ainda podem ser produzidos. Quando as lesões são unilaterais, tanto as aves quanto os pacientes humanos frequentemente recuperam algum comportamento vocal, porque o hemisfério oposto parece assumir parte da função afetada; além disso, a recuperação é melhor quando o canário é jovem ou quando o paciente é uma criança. Lesões parciais na estrutura NLC [outro núcleo cerebral vocal] dos papagaios causam déficits parciais na produção da estrutura acústica correta das vocalizações aprendidas, incluindo aquelas que correspondem às “falas”. Os sintomas são semelhantes aos da disartria em seres humanos que sofrem lesões em determinadas áreas do córtex motor facial. Lesões no córtex motor facial de chimpanzés e outros primatas não humanos não afetam sua habilidade de produzir vocalizações (JARVIS, 2004: 760-761)¹⁹.

Note, entretanto, que as eventuais homologias anatômicas e funcionais que o cérebro humano e os cérebros destes grupos de aves apresentam entre si com relação à capacidade de vocalização dizem respeito ao sistema sensorio-motor responsável pelo controle e pela programação dos movimentos que instanciam os sons da fala e do canto, respectivamente, e não fazem referência necessariamente a um núcleo propriamente linguístico. De fato, lesões na região do córtex motor primário, responsável pelo controle voluntário dos movimentos da face em seres humanos (cf. Fig. 2.6), fazem com que a geração de impulsos nervosos enviados aos músculos faciais seja interrompida ou não ocorra adequadamente, resultando em déficits motores que afetam o desempenho linguístico (como ocorre nas dispraxias ou disartrias), sem, contudo, atingir a competência gramatical.

¹⁹ “Lesions to songbird HVC and RA, particularly those placed on the left side in canaries, cause deficits similar to those found after damage to human face motor cortex, this being muteness for learned vocalizations, i.e., for speech. Innate sounds, such as crying and screaming, can still be produced, however. When the lesions are unilateral, both birds and patients often recover some vocal behavior, because the opposite hemisphere appears to take over some function; likewise, recovery is better when the canary is a juvenile or the patient a child. Partial lesions to parrot NLC causes partial deficits in producing the correct acoustic structure of learned vocalizations, including for learned speech. The symptoms are similar to that of dysarthria in humans after recovery from damage to the face motor cortex. Lesions to the face motor cortex in chimpanzees and other non-humans primates do not affect their ability to produce vocalizations.”

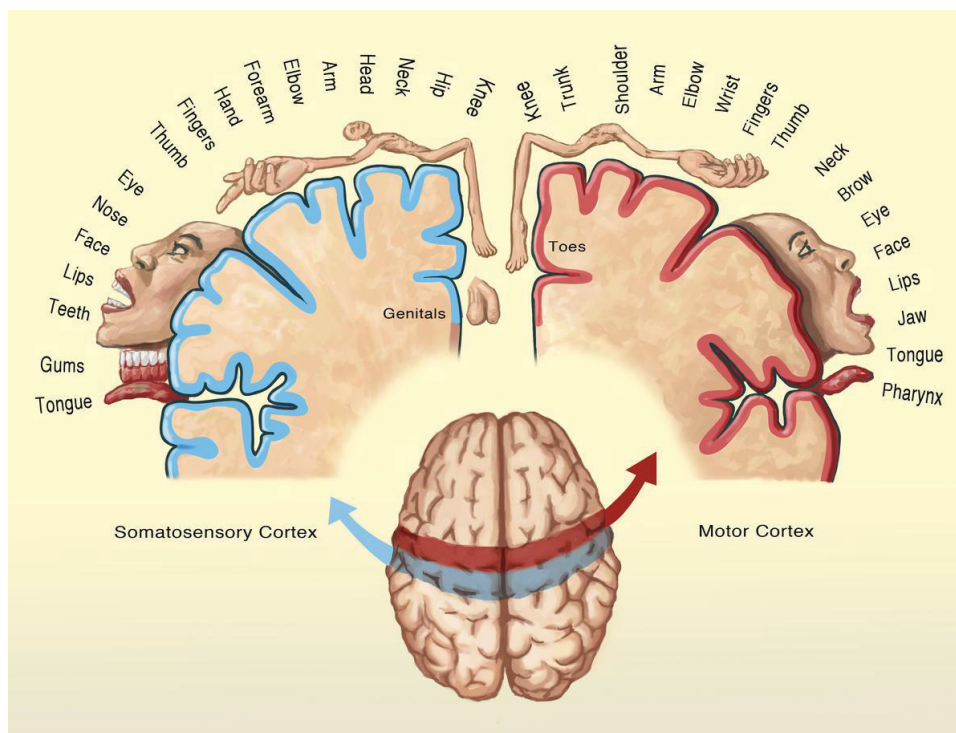


FIGURA 2.6. O homúnculo de Penfield (ou homúnculo cortical) é uma representação esquemática das áreas controladas pelo córtex motor (homúnculo motor) e pelo córtex somatossensorial (homúnculo sensorial). O tamanho de cada área do córtex é proporcional ao grau de controle motor ou ao grau de sensibilidade do órgão ou membro comandado; por isso, mãos, língua e rosto, que necessitam de um controle fino e são extremamente sensíveis, possuem contrapartes corticais relativamente extensas. Fonte da imagem: <https://www.sciencesource.com/archive/Cortical-Homunculus-Illustration-SS21056252.html#/SearchResult&ITEMID=SS21056252&DocRID=2OPEBM6P4488>. Acesso em 25/01/2021.

Nas apraxias e dispraxias da fala, as alterações/lesões atingem o planejamento motor e o indivíduo não consegue programar e executar os movimentos articulatórios necessários à produção adequada dos fonemas – de onde podem emergir alterações prosódicas secundárias (como pausas que “quebram” o fluxo natural do discurso e/ou fala “arrastada”). Nas disartrias, lesões no sistema nervoso central e/ou periférico acabam afetando a função motora de um dos (ou de vários) componentes musculares envolvidos na produção da fala (diafragma, cordas vocais, língua, lábios etc.), resultando em problemas de emissão (voz rouca ou sussurrada), de ritmo (fala excessivamente lenta ou excessivamente veloz/“atropelada”), entre outros (MURDOCH *et al.*, 1990). Nenhuma destas disfunções motoras, entretanto, é um reflexo necessário de uma disfunção propriamente linguística. Em indivíduos dispráxicos “puros” (i.e., naqueles em que a dispraxia não está associada a outros distúrbios neurológicos), por exemplo, as capacidades de leitura, de escrita e de compreensão da fala se mantêm intactas, indicando que a disfunção não alcançou o núcleo linguístico. Fatos como este apontam para a existência de uma arquitetura cognitiva em que a representação fonética ou gestual de uma

expressão linguística gerada pelo sistema gramatical é acessada por um sistema de interface – o sistema sensorio-motor (ou articulatório-perceptual) – que se responsabiliza pela computação dos movimentos musculares e pela articulação dos sons ou dos sinais visuais que materializam essa expressão linguística (Fig. 2.7). Se este sistema de desempenho é lesionado, os *enunciados* das expressões linguísticas – mas não elas mesmas – são atingidos.

Dada a arquitetura da linguagem que estamos assumindo neste ponto – que é aquela proposta por Chomsky (1995) e que obviamente está longe de ser a única –, além de uma representação fonética (π), que indica qual é a *forma* (sonora ou visual) que determinada construção gramatical deve ter, o sistema linguístico também deve associar a cada expressão produzida uma representação lógica (λ), que delimita o seu *sentido* e que é acessada pelo sistema de pensamento (conceitual-intencional) a fim de ser interpretada. Assim, cada expressão gerada pelo sistema computacional linguístico corresponde a um par do tipo *forma* – *significado* (π, λ) que pode ser acessado e processado por uma série de interfaces cognitivas externas ao sistema gramatical – dentre elas, o sistema articulatório-perceptual (responsável pela coordenação dos sons e dos gestos), que acessa (π), e o sistema conceitual-intencional (responsável pela construção dos sentidos), que acessa (λ).

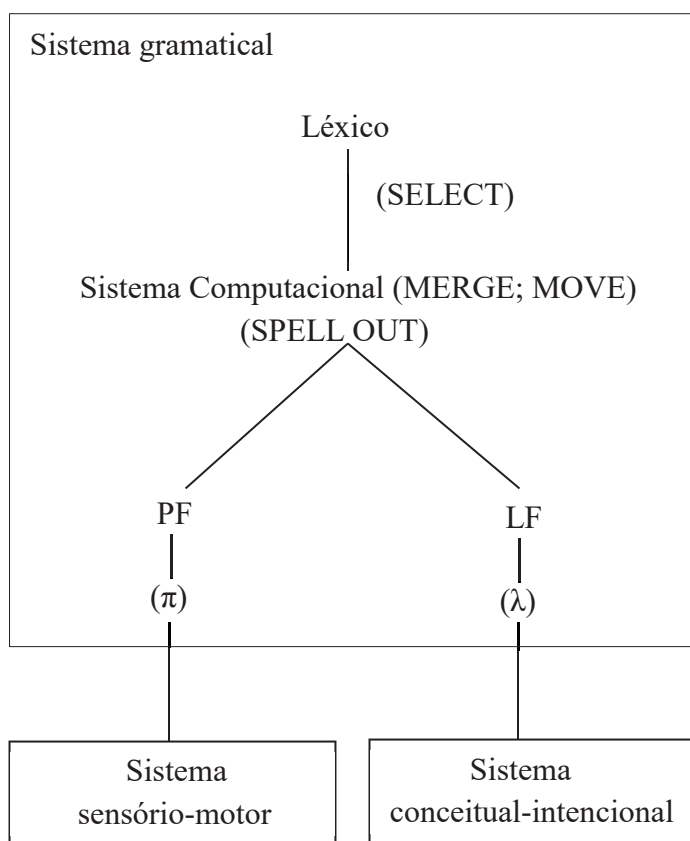


FIGURA 2.7. Uma versão simplificada da arquitetura da linguagem (cf. CHOMSKY, 1995: Cap. 4; KENEDY, 2013: Cap. 6). A derivação é o processo por meio do qual o sistema gramatical humano gera as representações linguísticas. Ela começa com a seleção (SELECT) dos elementos lexicais (palavras ou morfemas) que devem compor uma sentença; em seguida, o Sistema Computacional (Sintaxe) (i) combina recursivamente esses elementos através da operação MERGE, gerando sintagmas cada mais complexos, e (ii) promove o deslocamento (MOVE) de constituintes dentro da representação que está sendo construída (uma operação que consiste em um merge interno (ou *remerge*)). Finalmente, SPELL OUT separa as informações que serão enviadas para PF (*Phonetic Form*) e LF (*Logical Form*). “PF: componente da linguagem que converte as representações advindas da Sintaxe em instruções cognitivas (π) a serem enviadas para a interface sensorio-motora. LF: componente da linguagem que converte as representações advindas da Sintaxe em instruções cognitivas (λ) a serem enviadas para a interface conceitual-intencional” (KENEDY, 2013: 126).

Apesar de ser um modelo cognitivo, a arquitetura da linguagem apresentada na Fig. 2.7 está conceitualmente vinculada ao substrato neurológico que a realiza materialmente. Neste ponto, vale a pena ressaltar um pouco da complexidade da base neurofisiológica que dá sustentação à linguagem:

O próprio cérebro é um sistema muito complexo e a descrição da relação entre as funções linguísticas e o cérebro continua sendo um grande desafio. O cérebro consiste em matéria cinzenta e em matéria branca. A matéria cinzenta é composta por cerca de 100 bilhões de células neuronais [neurônios] que estão interconectadas por meio de trilhões de sinapses. Cada neurônio tem uma série de conexões (dendritos) por meio das quais *recebe* sinais de outros neurônios e também tem conexões (axônios) por meio das quais *envia* sinais a outros neurônios. Os axônios entram em contato com outros neurônios por meio de sinapses nas quais a transmissão dos sinais é realizada por neurotransmissores. A matéria branca, por sua vez, contém uma quantidade relativamente pequena de células neuronais, sendo composta por feixes de fibras que (i) conectam regiões cerebrais adjacentes por meio de feixes de fibras de curto alcance, ou (ii) conectam partes mais distantes do cérebro por feixes de fibras de longo alcance que garantem a comunicação entre elas. Em seu estado maduro, esses feixes de fibras são circundados por mielina, que serve como isolamento e permite a rápida propagação do sinal. Tanto a substância cinzenta quanto a branca são a base de todas as habilidades cognitivas, incluindo a linguagem (FRIEDERICI, 2017: 5)²⁰.

Como podemos notar pela complexidade desse mar organizado de neurônios, mapear o processamento neurológico da linguagem é como cartografar um labirinto (o da linguagem) sobre outro (o do cérebro), o que sem dúvida representa um desafio gigantesco. Mas se projetássemos a arquitetura da linguagem da Figura 2.7 no cérebro humano, poderíamos considerar que tanto as regiões corticais destacadas em vermelho (córtex motor primário) e em azul (córtex somatossensorial) na Figura 2.6 quanto as áreas de “aprendizagem vocal”

²⁰ “The brain itself is a very complex system, and the description of the relation between language functions and the brain remains a big challenge. The brain consists of gray matter and white matter. The gray matter is composed of about 100 billion neuronal cells that are interconnected via trillions of synapses. Each neuron has a number of connections via which it receives signals from other neurons (these are the dendrites), and it also has connections via which it forward signals to other neurons (these are the axons). The axons come into contact with other neurons via synapses at which the transmission of the signals is realized by neurotransmitters. The white matter, in contrast, contains only few neuronal cells, being composed of fiber bundles connecting adjacent brain regions by short-range fiber bundles, or connecting more distant parts of the brain by long-range fiber bundles that guarantee communication between these. In their mature state, these fiber bundles are surrounded by myelin, which serves as insulation and enables rapid propagation of the signal. Both the gray and the white matter are the basis for all cognitive abilities, including language.”

identificadas por Jarvis (2004) em humanos (que estão aproximadamente circunscritas ao córtex motor primário e ao córtex pré-motor, incluindo regiões como, por exemplo, a área de Broca) fazem parte do sistema sensório-motor²¹. As semelhanças entre os cérebros de seres humanos e de aves como papagaios, beija-flores e pássaros canoros, portanto, dizem respeito aos núcleos de controle motor da fala e do canto, mas não contemplam necessariamente os sistemas que geram abstratamente os produtos (linguísticos ou “canoros”) que serão acessados pelas interfaces articulatórias destas espécies.

Em estudo de 2014, Jarvis e sua equipe, analisando padrões de expressão gênica no tecido cerebral de mandarins-diamante, identificaram um grupo de 40 genes ativos ligados ao controle das vocalizações que são semelhantes aos que se expressam em áreas de processamento vocal no cérebro humano (PFENNING *et al.*, 2014). Sem dúvida, trata-se de

²¹ A Área de Broca, localizada em parte na região lateral esquerda do córtex pré-frontal e em parte na região pré-motora (cf. o *Apêndice*, relativo ao mapa neuroanatômico da linguagem), além de estar associada à programação motora, também é responsável, ao que tudo indica, por uma parcela do processamento gramatical. Tendo como base os resultados de um estudo experimental de observação do funcionamento cerebral durante o processamento de construções agramaticais (mais especificamente, de sequências que continham erros na ordem de um subgrupo de palavras, tais como **John drove to store the in a very fast car two weeks ago*) e durante o processamento de sentenças com erros de enunciação/ortográficos (tais como *John drove to the store in a very fasvt car two weeks ago*), Embick *et al.* (2000) propõem uma correlação entre a Área de Broca e o processamento sintático, dada a ativação neuronal significativa que essa área apresenta durante o processamento de agramaticalidades. Grodzinsky (2000), analisando um conjunto de dados referentes à afasia de Broca, atribui a esta área um papel bem mais restrito no processamento sintático. A Área de Broca, segundo o autor, seria responsável (i) pela computação das relações existentes entre os constituintes deslocados por regras transformacionais e os locais de extração (vestígios) deixados por eles na estrutura e (ii) pela construção das partes mais altas da estrutura de constituintes (a partir do nóculo temporal, mais especificamente). Na afasia de Broca, portanto, haveria (supostamente) (i) a deleção de todos os vestígios deixados pelos constituintes movidos pelas regras transformacionais, o que afetaria de modo significativo o processamento dos papéis temáticos, visto que a função- θ atribuída a um constituinte deslocado ocorre por via indireta, através da marcação temática de seu vestígio e (ii) déficits específicos relacionados à estruturação da sentença a partir do nó temporal, incluindo “problemas em partes ainda mais altas da árvore (CP). Como resultado, as perguntas-Wh e as orações encaixadas são inexistentes ou completamente mal formadas na fala dos pacientes. [...] Os dados atualmente disponíveis levam à conclusão de que pacientes afásicos agramáticos produzem árvores que estão intactas até o nó temporal [Tense node] e ‘podadas’ a partir desse nó [...]” (GRODZINSKY, 2000: 16). (Texto original: “[The impairment is also associated with] problems in still higher parts of the tree (CP). As a result, Wh-questions and embedded clauses are nonexistent or completely ill-formed in the speech of the patients. [...] The currently available data lead to the view that agrammatic aphasic patients produce trees that are intact up to the Tense node and ‘pruned’ from this node and up [...]”).

um achado fabuloso. Sem o controle motriz, cantos e falas não podem ser articulados e implementados. Note, entretanto, que estamos lidando com os genes que regulam os núcleos cerebrais responsáveis pela programação biomecânica das vocalizações e não com os genes que determinam o sistema (cerebral/mental) que as gera abstratamente, numa fase mais profunda da derivação destes produtos. Estamos, mais uma vez, na superfície da dimensão neurológica do uso concreto da fala e do canto, e não em sua dimensão mais abstrata, onde as expressões são geradas. Mas é exatamente nesta região profunda de processamento e geração de construções formais que estamos interessados. Se assumirmos, como os estudos aqui citados assumem, que seres humanos e aves como papagaios possuem, além de aparelhos fonadores, sistemas sensório-motores de controle neurológico da articulação sonora relativamente parecidos (inclusive do ponto de vista genético), por que existem diferenças tão significativas entre o que cada uma dessas espécies faz com esses dispositivos²²? Uma vez

²²Apesar de algumas semelhanças, as estruturas acústicas dos sons que podem ser produzidos por seres humanos e por papagaios são qualitativamente distintas – como também o são os dispositivos vocais acionados em sua produção. O controle voluntário do trato vocal supralaríngeo, por exemplo, parece ser uma característica exclusiva de nossa espécie (assim como a capacidade de produzir e encadear sons que não estão ligados obrigatoriamente a estados emocionais ou afetivos. Um cão emite sons com traços especiais quando está faminto ou quando vê comida (latidos e “gemidos” que são normalmente reconhecidos pelos tutores), mas nós podemos articular a sequência sonora *panqueca* sem estarmos com fome): “A regulação neural desses diferentes componentes fisiológicos da produção da fala parece envolver diferentes estruturas do cérebro, conectadas por vias complexas integradas em circuitos funcionais [...]. O controle voluntário do trato vocal supralaríngeo [SVT: *supralaryngeal vocal tract*] para fins linguísticos parece ser uma característica humana única. As estruturas neurais que estão implicadas na regulação do SVT humano essencialmente armazenam e acessam rapidamente padrões automatizados de atividade muscular que têm uma função linguística e não emotiva. Comandos musculares automatizados são, em essência, respostas “superaprendidas” que são executadas sem pensamento ou esforço conscientes. Os padrões automatizados do SVT da fala humana são amplamente independentes do afeto e do estado psicológico. Os seres humanos, portanto, podem produzir sons que têm uma relação arbitrária com seu estado emocional. Os sons que compõem a palavra *help*, por exemplo, não possuem em si nenhuma qualidade emocional inerente. O som [h] também pode ser usado na palavra *hello*. Em contraste, os sinais vocais de outros mamíferos parecem estar ligados ao seu estado afetivo. Um chimpanzé produzirá uma vocalização específica ao ver comida. No entanto, ele não pode produzir uma sequência arbitrária de sons para sinalizar *food*.” (LIEBERMAN, 1996: 57). (Texto original: “The neural regulation of these different physiologic components of speech production appears to involve different structures of the brain, connected by complex pathways into functional circuits [...]. The voluntary control of the supralaryngeal vocal tract for linguistic ends appears to be a unique human characteristic. The neural structures that are implicated in the regulation of the human SVT essentially store and rapidly access automatized patterns of muscular activity that have a linguistic rather than an emotive function. Automatized muscular commands are,

que são aptos bioacusticamente, será que não falta aos papagaios (e aos outros “seres sonoros”) um sistema cognitivo especializado que capacite essas aves a usar todo esse aparato vocal do modo como nós, humanos, usamos? Ou seja, será que não falta a esses animais um sistema propriamente *linguístico*?

Trabalhos como o de Irene Pepperberg (1990, 1992, 1999), que desde a década de 70 desenvolve pesquisas sobre cognição animal com papagaios-cinzentos (*Psittacus erithacus*), têm demonstrado sistematicamente que, apesar dos avanços significativos em direção a uma melhor compreensão da forma como estes animais “pensam” e se comunicam, existem barreiras linguísticas que parecem intransponíveis quando se analisam as habilidades comunicativas alcançadas por eles, indicando que, sem um módulo cognitivo específico que oriente o aprendizado, a aquisição de uma língua se torna impossível. A principal destas barreiras – que é ao mesmo tempo uma função biológica e linguística – é a recursividade. Durante a derivação de uma expressão, o sistema computacional humano é capaz de aplicar recursivamente a mesma regra de combinação de elementos (Merge), o que permite a geração (potencial) de infinitas construções linguísticas a partir de um conjunto finito de recursos. Para Hauser, Chomsky e Fitch (2002), aliás, a recursividade é a operação linguística por excelência e representa a propriedade central da faculdade da linguagem:

Todas as abordagens concordam que uma propriedade central da faculdade da linguagem em sentido estrito [FLN: *faculty of language – narrow sense*] é a recursão, atribuída à sintaxe na concepção que acabamos de delinear [i.e., a sintaxe como um sistema computacional]. FLN pega um conjunto finito de elementos e produz uma gama de expressões discretas potencialmente infinita. É esta capacidade da FLN que produz infinitude discreta (uma propriedade que também caracteriza os números naturais). (HAUSER *et al.*, 2002: 1571)²³.

in essence, ‘overlearned’ responses that are performed without conscious thought or effort. The automatized SVT patterns of human speech are largely independent of affect and psychological state. Humans, therefore, can produce sounds that have an arbitrary relationship to their emotional state. The sounds that make up the word *help*, for example, in themselves do not have any inherent emotional quality. The sound [h] can just as well be used in the word *hello*. In contrast, the vocal signals of other mammals appear to be tied to their affective state. A chimpanzee will produce a food-bark on seeing food. However, he cannot produce an arbitrary sequence of sounds to signal *food*.”).

²³ “All approaches agree that a core property of FLN [Faculty of language – narrow sense] is recursion, attributed to narrow syntax in the conception just outlined. FLN takes a finite set of elements and yields a potentially infinite array of discrete expressions. This capacity of FLN yields discrete infinity (a property that also characterizes the natural numbers).”

Segundo os autores, é a recursividade o mecanismo que nos singulariza linguisticamente com relação ao restante do reino animal. Sendo ou não uma ferramenta cognitiva exclusiva de nossa espécie, o fato é que nem mesmo Alex, o mais brilhante dos papagaios-cinzentos treinados por Pepperberg, demonstrou ter alcançado um domínio satisfatório da recursão e desta capacidade imanente de gerar infinitas expressões. Alex era capaz de discriminar cores, formas e tipos de material (papel, madeira, couro etc.) e de combinar corretamente estes traços durante tarefas de rotulagem de objetos e de seus atributos (*labeling*), produzindo construções do tipo “*rose paper*” ou “*green wood*”. Habilidades como estas indicam um grau considerável de representação mental e de manipulação de conceitos abstratos, além de apontarem para a existência de princípios de categorização que permitem a formação de classes de objetos. “*Alex realmente aprendeu a formar classes categóricas. Ele podia (i) compreender uma pergunta oral, (ii) extrair a categoria relevante da pergunta e de um objeto que poderia ser classificado de várias maneiras, e (iii) responder com o rótulo adequado, nomeando a única instância correta dessa categoria.*” (PEPPERBERG, 1999: 61)²⁴. Assim, as respostas (na maioria das vezes, corretas) de Alex a perguntas do tipo “*What color?*” ou “*What shape?*” indicavam que ele talvez fosse capaz de operar com relações hierarquizadas envolvendo instâncias particulares (hipônimos: *yellow, blue, green*) de uma categoria (hiperônimo: *COLOR*).

Papagaios podem, através da imitação, adquirir um repertório mais ou menos extenso de cadeias fonéticas extraídas de construções linguísticas (arranjos sonoros que vão desde “palavras soltas” até trechos de canções), memorizando, inclusive, o contexto em que algumas delas são utilizadas (como “bom dia”, “boa noite” etc.), sem, claro, atentar para os significados de suas partes ou mesmo o seu sentido global. Expressões simples (como (i) *o falcão pegou o papagaio*) podem nos dar a impressão de que o significado global de uma construção linguística deriva simplesmente da soma linear dos significados dos itens lexicais que a compõem. No entanto, expressões mais complexas (como (ii) *o falcão pegou o papagaio com as garras afiadas*) mostram que o sentido de uma construção deriva não apenas dos sentidos das suas partes, mas também da forma como essas partes estão combinadas e estruturadas (*Princípio de Composicionalidade* (FREGE, 1978 [1892])) – o que permite, entre outras coisas, que uma mesma sequência de palavras possa representar múltiplas estruturas

²⁴ “Alex had indeed learned to form categorical classes. He could comprehend a vocal question, extract the relevant category from the question and from an object that could be classified in multiple ways, and respond with the label for the one, correct instance of this category.”

distintas, cada uma delas com um sentido próprio. Essa ambiguidade estrutural é o que faz com que a expressão (ii) possa representar duas construções que, apesar de possuírem formas fonéticas (quase que totalmente) coincidentes, são estrutural e semanticamente distintas:

(ii') O falcão [pegou [o papagaio com as garras afiadas]] (= foi o papagaio com as garras afiadas que o falcão pegou).

(ii'') O falcão [[pegou [o papagaio]] [com as garras afiadas]] (= foi com as garras afiadas que o falcão pegou o papagaio).

Apesar de eventualmente conseguirem reproduzir “biomecanicamente” as cadeias fonéticas de construções como (ii), papagaios (e, ao que tudo indica, todo o restante do reino animal) não conseguem acessar as estruturas sintático-semânticas subjacentes a estas construções. Alex, por exemplo, de tanto ouvir as questões formuladas pelos treinadores, adquiriu os desenhos melódicos de algumas perguntas, sem, no entanto, demonstrar que as usava de modo intencional quando as pronunciava (com, talvez, uma única exceção²⁵). Aparentemente, ele era incapaz de processar semanticamente a maioria das expressões que memorizava.

E mesmo quando Alex demonstrava *alguma* criatividade linguística, os produtos referencialmente ativos de sua competência combinatória (ou seja, aquelas expressões vocais que possuíam um conteúdo semântico e referencial claramente identificáveis e que não eram simplesmente manifestações miméticas) dificilmente apresentavam mais de dois termos/conceitos associados (por exemplo, *blue key*), indicando uma limitação do seu sistema

²⁵ “O uso produtivo de Alex das frases ‘*What color?*’ e ‘*What shape?*’, que seria outra medida de capacidade cognitiva, ainda era limitado. Na época, ele havia usado a frase ‘*What color?*’ apenas uma vez de uma maneira que poderia ser interpretada como intencional: ele perguntou a um aluno sobre a cor de sua imagem no espelho [...]. Depois de ouvir respostas como ‘*That’s grey; you’re a Grey parrot*’ para seis de suas consultas, ele adquiriu o rótulo *grey* e posteriormente o usou para identificar corretamente, logo nas primeiras exposições, objetos tingidos de cinza.” (PEPPERBERG, 1999: 61. Cf. o texto original ao final desta nota). Talvez este momento de “autorreconhecimento” de Alex tenha sido um mero acidente. Não deixa, entretanto, de ser um episódio que mereça atenção. Superestimar as habilidades de uma espécie pode induzir ao erro; subestimá-las pode representar uma barreira para a compreensão de suas reais capacidades. (Texto original: “Alex’s productive use of the phrases ‘*What color?*’ and ‘*What shape?*’, which would be another measure of cognitive ability, was still limited. At the time, he had used the phrase ‘*What color?*’ only once in a way that could possibly be interpreted as intentional: He queried a student as to the color of his mirror image [...]. After hearing replies such as ‘*That’s grey; you’re a Grey parrot*’ to six of his queries, he acquired the label *grey* and subsequently used it to identify correctly, on first showings, newly dyed grey objects.”).

de combinação. Com base nos dados apresentados por Pepperberg sobre as habilidades identificadas em Alex e em outros indivíduos da mesma espécie, podemos eventualmente desconfiar que o sistema computacional/gramatical dos papagaios-cinzentos possui algum tipo de princípio combinatório parecido com Merge (já que é capaz de combinar dois constituintes, A e B, formando um terceiro, C (= AB), estruturalmente mais complexo: $A + B \rightarrow C (= AB)$; *want + banana* $\rightarrow$ *want banana* [PEPPERBERG, 1999: 205]), mas que não pode ser aplicado recursivamente de modo a produzir uma derivação do tipo: Merge 1: {aquela, banana}; M2: {quero, {aquela, banana}}; M3: {eu, {quero, {aquela, banana}}}. Assim, as saídas (*outputs*) binárias geradas pelo primeiro Merge aparentemente não podem ser utilizadas como entradas (*inputs*) de uma nova combinação, o que impede a derivação de estruturas progressivamente mais complexas. Sem que possa ser aplicada recursivamente, uma operação gramatical de combinação mínima (classifiquemos ela como Merge ou não) gera apenas binômios e não infinitude discreta – ou, sob outra perspectiva, podemos admitir que há alguma recursão (especialmente no encadeamento recursivo de sons que determinadas aves canoras empregam na produção de seus cantos naturais), mas não *recursão ilimitada* (BERWICK, 2011).

É preciso lembrar, entretanto, que os estudos de Pepperberg partem de um quadro teórico bem distinto do que estamos discutindo aqui. Suas pesquisas sobre o desenvolvimento de habilidades comunicativas em aves têm como base as ideias de psicólogos como Jean Piaget, Lev Vygotsky e, principalmente, a Teoria da Modelagem Social (BANDURA, 1977), de Albert Bandura. Para estes autores, a aquisição de uma língua, assim como de qualquer tipo de conhecimento, depende de mecanismos gerais de aprendizagem e não de um módulo linguístico específico cuja operação central é a recursão:

A Teoria da Modelagem Social, portanto, enfatiza como a atenção, compreensão e motivação afetam a aprendizagem. A teoria consiste em um conjunto de princípios [...] que descrevem a forma ótima do *input* social para qualquer tipo de aprendizagem [...]. Como alguns princípios presumem a existência de habilidades cognitivas (como codificação simbólica, organização cognitiva e representação mental) que não podem ser atribuídas inquestionavelmente a todos os não humanos [...], nem todos os princípios podem ser relevantes para os estudos com animais. Acredito, entretanto, que um subconjunto de quatro princípios é claramente aplicável à aprendizagem vocal em aves [...]. Um princípio afirma que o nível de competência do aluno deve ser levado em consideração [...]. Assim, crianças geralmente adquirem nomes de cores adicionais mais rapidamente do que os primeiros nomes de cores. Alguns pássaros canoros parecem responder de forma semelhante, no sentido de que adquirem mais facilmente vocalizações aloespecíficas que se assemelham a seus cantos naturais, no que se refere, por

exemplo, ao padrão tonal ou à sintaxe. [...] Um segundo princípio da Teoria da Modelagem Social afirma que a modelagem deve ajudar o aprendiz a entender como o novo material se relaciona com os problemas atuais e que vantagem é conferida pelo aprendizado do novo material (novamente, observe os paralelos com Piaget e Vygotsky). [...] Assim, por exemplo, uma criança a quem é mostrado exatamente como pedir um brinquedo, ou um pássaro que vê exatamente como uma determinada vocalização é usada em um encontro agressivo, provavelmente aprenderão mais prontamente do que se estiverem em situações sem tais demonstrações. [...] Um terceiro princípio afirma que quanto mais intensa é a interação entre um aluno e seus modelos, mais eficaz é o treinamento. [...] O quarto princípio afirma que se existe inibição ou resistência em relação ao aprendizado, os três primeiros princípios são ainda mais importantes (PEPPERBERG, 1999: 21-24)²⁶.

Para Pepperberg, todos estes princípios devem compor um instrumental que permita a aquisição, pelo aprendiz, de três componentes que, em conjunto, formam o *minimum minimorum* para que uma comunicação seja realmente efetiva: “(1) a semântica do código (o significado de seus elementos individuais), (2) sua sintaxe (as regras apropriadas para combinar seus elementos) e (3) sua pragmática (como o código é usado e como o seu uso afeta outros indivíduos).” (PEPPERBERG, 1999: 20)²⁷.

²⁶ “Social modeling theory thus emphasizes how attention, comprehension, and motivation affect learning. The theory consists of a set of principles [...] that describe the optimal form of social input for any type of learning [...]. Because some principles presume the existence of cognitive abilities (such as symbolic coding, cognitive organization, and mental rehearsal) that cannot be unquestionably attributed to all nonhumans [...], not all principles may be relevant for animal studies. I believe, however, that a subset of four principles is clearly applicable to avian vocal learning [...]. One principle states that the student’s level of competence must be taken into account. [...] Thus children generally acquire additional color labels more quickly than their first color labels. Some songbirds seem to respond similarly, in that they more easily acquire allospecific vocalizations that resemble their natural songs in, for example, tonal pattern or syntax. [...] A second principle of social modeling theory states that modeling must help the student understand how new material relates to current problems and what advantage is conferred by learning new material (again, note the parallels with Piaget and Vygotsky). [...] Thus, for example, a child who is shown exactly how to request a toy, or a bird that sees exactly how a particular vocalization is used in an aggressive encounter, are both likely to learn more readily than if they were in situations without such demonstrations. [...] A third principle states that the more intense the interaction between a student and its models, the more effective is the training. [...] The fourth principle states that if inhibition or resistance exists toward learning, the first three principles are even more important.”

²⁷ “[I have previously argued that for true communication to develop, tutoring must expose a student to three components of the desired communication code]: (1) the code’s semantics (the meaning of its individual elements), (2) its syntax (the appropriate rules for combining its elements), and (3) its pragmatics (how the code is used and how its use affects other individuals).”

Note que os princípios adotados por Pepperberg são orientados por um modelo piagetiano de desenvolvimento cognitivo, na medida em que a aprendizagem (dos nomes das cores pelas crianças ou das vocalizações aloespecíficas pelas aves, por exemplo) é concebida como um processo que envolve a construção subjetiva de conhecimentos e a assimilação gradativa de objetos cada vez mais complexos – objetos estes que, ao serem assimilados/aprendidos, modificam as próprias estruturas cognitivas que os representam, tornando-as mais poderosas e capazes de manipular objetos da mesma natureza. Assim, a assimilação de um objeto que, pelo seu grau de complexidade, não consegue ser representado e processado adequadamente pelo esquema cognitivo já existente no indivíduo, faz com que esse esquema seja reconfigurado a fim de se acomodar à complexidade dos novos dados advindos da experiência (PIAGET, 1975, 1979). Por isso, objetos que fazem parte de classes já assimiladas – como, por exemplo, as vocalizações que são moduladas em um padrão tonal já conhecido pela ave – são mais facilmente adquiridos pelo indivíduo, uma vez que seu esquema cognitivo já se encontra ajustadamente formatado para receber a nova informação.

Modelos de desenvolvimento como o que Pepperberg assume em seu programa de pesquisa são baseados em mecanismos gerais de cognição que podem ser mobilizados para a execução de diferentes tarefas mentais. Assim, a aquisição, pelas aves, de um sistema formado por signos e por regras de construção e transmissão de mensagens não dependeria de um módulo especialmente configurado para este fim, mas de funções cognitivas polivalentes – como a capacidade de representação simbólica e a formação de memórias. De forma semelhante, a aquisição de uma língua, pela criança, não ocorreria graças a um módulo especificamente programado para processar os dados linguísticos ambientais e formatá-los sob a forma de uma gramática internalizada, mas decorreria de mecanismos intelectivos gerais de processamento da experiência e de construção de conhecimento²⁸.

Apesar de Alex ter adquirido um conjunto de itens lexicais relativamente extenso – incluindo algumas fórmulas verbais como *want X*, em que o complemento X, quase sempre,

²⁸ Para Jean Piaget, por exemplo, a aquisição de uma língua decorre daquilo que o autor chama de *função semiótica* – grosso modo, a capacidade que as crianças adquirem de evocar objetos ausentes e de representá-los em esquemas mentais. Assim, a linguagem se configura como a manifestação particular de uma função simbólica mais geral. Para uma discussão concisa sobre este conceito piagetiano, cf. Piatelli-Palmarini (1980): cap. VII.

referia-se a algum tipo de alimento (biscoito, nozes, banana)²⁹ – e de aplicá-lo com certa engenhosidade em tarefas de nomeação e de requisição (cf. o cap. XI de Pepperberg (1999), em que a autora discute a (provável) intencionalidade dos pedidos de Alex), não é possível estabelecer com exatidão até que ponto a competência comunicativa dos papagaios-cinzentos dimana de uma verdadeira (proto)gramática. Existem vestígios de processamento linguístico nos fragmentos de falas produzidos por alguns destes animais, mas eles são tão frágeis e tão fugidios que acabam realçando, por contraste, o imenso abismo gramatical que nos distancia.

2.2.3. Primatas “sinalizadores”

A homologia molecular existente entre o DNA da espécie humana e o de chimpanzés (gênero *Pan*) e gorilas (gênero *Gorilla*) é de aproximadamente 99% e 98%, respectivamente

²⁹ Sobre a possível criatividade linguística de Alex, Pepperberg se mostra (relativamente) cética: “A substituição lexical requer muito menos conhecimento do código de comunicação do que a criatividade lexical [...]; o primeiro mecanismo envolve apenas a substituição de um elemento por um item relacionado em uma construção estocada [armazenada]. Um exemplo é o comportamento de um indivíduo que, tendo aprendido ‘*Polly quer biscoito*’, subsequentemente aprende ‘*bolacha*’ e produz espontaneamente ‘*Polly quer bolacha*’ [...]. O indivíduo meramente segmentou a sentença original em “objetos desejados” (biscoito, etc.) e uma construção estocada, ‘*Polly quer...*’. A substituição lexical não é uma operação trivial porque requer conhecimento semântico e conhecimento sintático limitado [...]; isto é, um indivíduo deve, por exemplo, saber substituir um nome por um nome e manter a ordem das palavras. Um indivíduo capaz de substituição lexical, entretanto, não é necessariamente capaz de criatividade lexical. A substituição lexical é uma habilidade passível de ser ensinada a primatas não humanos [...], mamíferos marinhos [...] e humanos com deficiência [...]. Com relação a duas construções estocadas, a substituição lexical de Alex pode, como a das crianças, ter sido incentivada por mapeamento referencial. Refiro-me à sua colocação espontânea de novos rótulos (*labels*) nos *slots* X e Y das frases de estoque ‘*I want (wanna) X*’ e ‘*Wanna go Y*’” (PEPPERBERG, 1999: 245). (Texto original: “Lexical substitution requires much less knowledge of the communication code than does lexical creativity [...]; the former merely involves substituting one element with a related item in a stock phrase. An example is the behavior of a subject who, having learned ‘*Polly want a cracker*’, subsequently learns ‘*cookie*’, and spontaneously produces ‘*Polly want a cookie*’ [...]. The subject has merely segmented the original sentence into “objects wanted” (cookie, etc.) and a stock phrase, ‘*Polly want ...*’. Lexical substitution is nontrivial because it requires semantic and limited syntactic knowledge [...]; that is, a subject must, for example, know to substitute a noun for a noun and to maintain word order. An individual capable of lexical substitution, however, is not necessarily capable of lexical creativity. Lexical substitution is an educable skill in nonhuman primates [...], marine mammals [...], and handicapped humans [...]. With respect to two stock phrases, Alex’s lexical substitution may, like those of children, have been encouraged by referential mapping. I refer to his spontaneous placement of new labels in the X and Y slots of the stock phrases ‘*I want (wanna) X*’ and ‘*Wanna go Y*’.”).

(Chimpanzee Sequencing and Analysis Consortium, 2005; SCALLY *et al.*, 2012). Isto significa que, em média, há apenas uma ou duas bases divergentes a cada arranjo de cem bases nitrogenadas que formam os programas bioquímicos de seres humanos e primatas superiores. Uma homologia tão robusta a nível molecular, é claro, reflete-se fenotipicamente em inúmeras características (físicas e cognitivo-comportamentais) que se assemelham entre essas espécies e a nossa – como a presença de polegares opositores, o bipedismo (relativo nos grande símios), a tendência a se organizar socialmente, etc. De fato, é no mínimo intrigante ver um gorila bebê mamando no colo de sua mãe e perceber os pontos de convergência de nossas naturezas e de nossas linhas evolutivas. Mas existem também dessemelhanças, e elas são significativas. Os grandes macacos possuem 24 pares de cromossomos homólogos ($2n = 48$), enquanto nós, seres humanos, possuímos 23 ($2n = 46$). O cromossomo humano número 2 está dividido em dois cromossomos mais curtos nos primatas superiores, indicando que, em algum momento da linha evolutiva humana, uma mutação provocou a fusão de dois cromossomos, diminuindo nosso lote cromossômico, rearranjando os genes e nos distanciando do ramo evolutivo que deu origem aos macacos modernos. Além disso, o 1% gênico exclusivo de nossa espécie interage de forma complexa com os outros 99% e com o ambiente, resultando em diferenças fenotípicas marcantes entre chimpanzés e gorilas, de um lado, e seres humanos, de outro. Homologias bioquímicas não se projetam, necessariamente, em fenótipos homólogos. Aproximadamente 60% dos nossos genes são semelhantes aos de moscas-das-frutas (MIRZOYAN *et al.*, 2019), por exemplo, mas isto não significa que a semelhança *fenotípica* existente entre humanos e drosófilas seja da ordem de 60%. As moscas voam e têm olhos compostos; os seres humanos constroem aviões e caleidoscópios. A semelhança *genotípica*, apesar de ser uma evidência fabulosa da evolução da vida e da relação entre as espécies, não reflete a semelhança fenotípica entre elas. Assim, apesar da imensa área de sobreposição gênica que pode ser mapeada entre o DNA de seres humanos e o de primatas superiores, as interações gênicas e proteicas que emergem do genoma de cada espécie criam configurações biológicas únicas, com capacidades únicas. Se há continuidade genética entre esses grupos – que, no limite, derivam de um ancestral comum –, há também descontinuidades fundamentais. Há um certo “vale da consciência” que nos distancia dos nossos parentes primatas e nos descola do reino animal – como se o *Homo sapiens sapiens*, em determinado momento de sua evolução, tivesse alcançado, como resultado de mutações genéticas acumuladas ao longo de centenas de milhares de anos, a singularidade da condição humana.

Primatas superiores são capazes de construir ferramentas, mas só nós somos capazes de aperfeiçoá-las ao longo das gerações, transmitindo conhecimento cultural acumulado através dos costumes, da fala, dos cantos, dos ritos e dos mitos; primatas têm lembranças e fazem planos para o futuro imediato (chimpanzés montam estratégias para caçar em grupo, por exemplo), mas só nós podemos falar sobre o passado profundo, quando ainda não existíamos, e sobre o tempo que ainda virá, quando não mais existiremos; se um gorila inadvertidamente mata um filhote de seu grupo, é pouco provável que sinta o mesmo abalo emocional e que sofra a mesma sanção moral que recairá sobre a consciência de um ser humano em uma situação semelhante (ao menos, sobre a consciência da maioria de nós), pois só nossa espécie possui a capacidade de interiorizar valores morais socialmente construídos; há registros melancólicos de chimpanzés fêmeas que continuam carregando e protegendo seus filhotes mesmo depois de mortos, mas não há angústia comparável à de uma mãe humana que vê o seu filho sendo enterrado – exatamente porque esta, diferentemente daquelas, sabe que aquele corpo, sob a terra, representa o fim; grandes símios formam grupos e entram em conflito por território, mas só nós construímos civilizações e usamos toda nossa capacidade estratégica para exterminar populações rivais; chimpanzés e gorilas se comunicam com desenvoltura e podem adquirir um vocabulário de sinais relativamente extenso quando treinados sistematicamente, mas só os seres humanos adquirem a língua da comunidade de fala (ou sígnica) em que estão inseridos sem esforço aparente e sem que sejam orientados por tutores. Estas e outras diferenças interespecíficas colocam em evidência os limites dentro dos quais a plasticidade fenotípica se realiza: cada cérebro, de cada espécie, deriva de um cluster específico de genes neuromodeladores e, por isso, apresenta uma configuração funcional específica. Assim como o cérebro humano não consegue processar instintivamente o intervalo de tempo existente entre a emissão, reflexão e recepção de ondas sonoras de modo a construir um mapa mental do ambiente em que se locomove – como o fazem golfinhos e morcegos durante a ecolocalização –, é de se esperar que algumas funções cognitivas sejam exclusivamente humanas – dentre elas, a competência linguística, i.e., a capacidade de produzir significados que possam ser veiculados através de cadeias formais (de sons, de sinais visuais, etc.)³⁰.

³⁰ Chomsky (2011) considera que tanto a “externalização” (a produção material de sequências linguísticas pelo sistema sensorio-motor) quanto a comunicação (a transmissão de mensagens) são propriedades secundárias da linguagem, cuja função primordial, segundo o autor, é a produção/construção de significados por meio de sons (ou outros meios, como sinais visuais). Por isso, Chomsky sugere que ocorra uma inversão da ideia aristotélica

Assumir que a faculdade da linguagem é uma dotação biológica exclusiva da espécie humana significa admitir que mesmo animais evolutivamente próximos (como chimpanzés, bonobos e gorilas) não podem, entre outras coisas, criar sentenças. Será? Será que a chimpanzé Washoe (GARDNER & GARDNER, 1969) e a gorila Koko (PATTERSON, 1978), que ficaram mundialmente conhecidas por suas habilidades comunicativas e por seu extenso repertório de vocábulos visuais adaptados da *American Sign Language* (ASL, a língua de sinais norte-americana), não eram capazes de organizar estes itens lexicais de forma criativa, construindo, através de regras de combinação, sintagmas e sentenças “inéditos” (i.e., não memorizados)?

Consideremos a seguinte construção:

(2) Aquele gorila comeu um bolo de banana caramelizada na praia.

Desde que conheça o significado dos itens lexicais que compõem a sentença (2), qualquer falante da língua portuguesa, mesmo que nunca tenha sido exposto a esta combinação única de elementos, é capaz de processá-la semanticamente em sua totalidade, atribuindo a ela (pelo menos) uma interpretação³¹. Usar criativamente uma língua significa

segundo a qual linguagem é “som com significado”: “Parece que a máxima de Aristóteles deveria ser invertida: a linguagem não é som com significado, mas significado com som, uma coisa muito diferente. A externalização pelo sistema SM [sistema sensorio-motor] parece ser uma propriedade secundária da linguagem. A externalização também é, em parte, independente da modalidade, como os trabalhos sobre as línguas de sinais têm revelado nas últimas décadas. Às vezes a externalização é empregada para a comunicação – mas nem sempre (pelo menos não se nós estamos investindo o termo “comunicação” de alguma significância). Assim, a comunicação, *a fortiori*, é uma propriedade ainda mais auxiliar da linguagem, ao contrário do que afirmam muitas doutrinas convencionais – e, claro, o uso da linguagem é apenas uma das muitas formas de comunicação.” (CHOMSKY, 2011: 275). (Texto original: “It appears that Aristotle’s dictum should be reversed: language is not sound with meaning but meaning with sound, a very different matter. Externalization by the SM system appears to be a secondary property of language. Externalization is also in part independent of modality, as work of the past few decades on sign language has revealed. Sometimes externalization is employed for communication – by no means always, at least if we invest the term “communication” with some significance. Hence, communication, *a fortiori*, is a still more ancillary property of language, contrary to much conventional doctrine – and of course language use is only one of many forms of communication.”).

³¹ Em (2), o sintagma preposicional [_{pp} *na praia*] pode fazer referência tanto ao local em que o gorila comeu o bolo (cf. a interpretação (i): *Foi na praia que aquele gorila comeu um bolo de banana caramelizada*) quanto ao local em que a fruta foi caramelizada (cf. a interpretação (ii): *Foi um bolo de banana caramelizada na praia que aquele gorila comeu*). Note que as duas interpretações disponíveis em (2) são mutuamente exclusivas e

exatamente isso: processar e interpretar arranjos inéditos de constituintes a partir de peças lexicais já conhecidas e de regras de combinação. Perceba que o conceito de *criatividade linguística* que estamos usando aqui se refere ao núcleo de processamento gramatical da linguagem e se relaciona tanto à produção quanto à recepção de construtos linguísticos; logo, se alguém é capaz de interpretar adequadamente a sentença (2), este alguém está fazendo um uso criativo da língua, na medida em que é capaz de reconstruir em sua mente uma estrutura de significado inédito. A infinitude discreta da qual somos dotados não diz respeito apenas às infinitas sentenças que podemos gerar “de dentro para fora” (ou “da mente para o mundo”), mas também contempla as infinitas construções linguísticas que nos chegam do mundo e que podemos gerar “de fora para dentro”.

Consideremos agora a sequência de quadros da Figura 2.8, em que a chimpanzé Washoe interage com sua tutora, Beatrix Gardner. Os quadros finais – de 9 a 12 – mostram que Washoe produz subsequentemente os sinais */time/* e */eat/* após ter sido exposta a cada um deles de forma separada (*/time/* no quadro 8; */eat/* no quadro 1). Será que esta sequência de sinais representa um rearranjo criativo de itens já conhecidos? Será que Washoe foi capaz de construir um sintagma sinalizado internamente complexo do tipo */time eat/* (em tradução livre: “hora de comer”), em resposta à pergunta sinalizada */What time (is it) now/*, feita por Beatrix Gardner?

correspondem a duas estruturas distintas: *Aquele gorila* [_{VP} [_{VP} *comeu um bolo de banana caramelizada*] [_{PP} *na praia*]] vs. *Aquele gorila comeu um bolo de* [_{DP} *banana caramelizada na praia*].

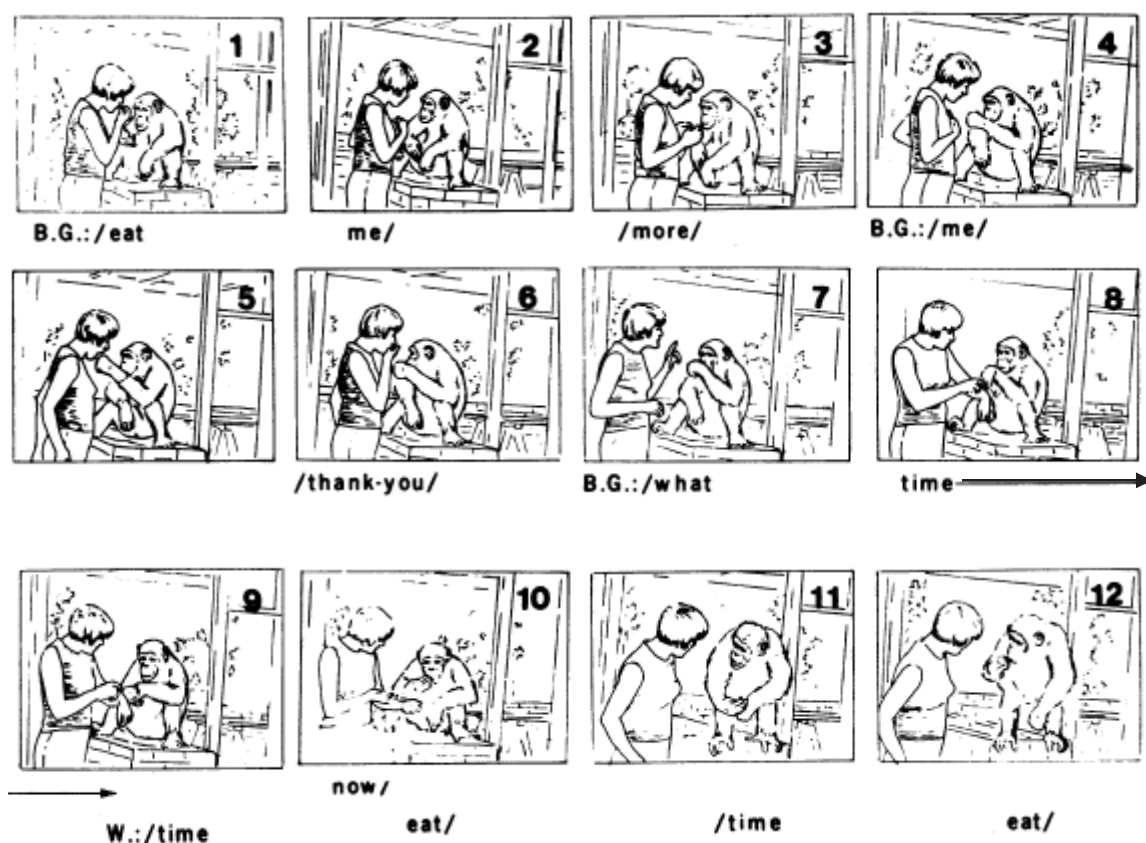


FIGURA 2.8. *Storyboard* (feito a partir de filmagem) representando uma sequência comunicativa entre Beatrix Gardner (B.G.) e a chimpanzé Washoe (W.). A seta que atravessa os quadros (8) e (9) representa a continuidade do mesmo sinal (*time*), feito por B.G. durante a construção do enunciado interrogativo */what time now/*. Os quadros (9) e (10) mostram que Washoe não espera a troca de turnos típica dos diálogos e começa a sinalizar durante o turno de B.G. (cf. nota 33). Extraído e adaptado de Terrace *et al.*, 1979.

Exemplos como o da Figura 2.8 mostram que, na prática, não é simples fazer a distinção entre o que é um arranjo consciente e criativo de dois itens lexicais que, combinados, formam um novo constituinte ($A + B \rightarrow C$, sendo $C = AB$) e o que é uma simples justaposição temporal de dois elementos que não estabelecem entre si nenhum tipo de relação gramatical interna ($A + B \rightarrow A + B$). No caso da interação comunicativa da Figura 2.8, trata-se de descobrir se o *link* existente entre */time/* e */eat/* é construído criativamente pela mente de Washoe ou se este *link* é “acidentalmente” construído pela competência gramatical de quem observa e processa este conjunto de expressões³². Analisando este e um conjunto

³² Estudos sobre cognição animal podem apresentar dois tipos de desvios com relação à metodologia empregada e à interpretação de seus resultados: o *Efeito Clever Hans* – quando o treinador, inadvertidamente, transmite dicas ao animal sobre a resposta adequada a uma determinada tarefa (olhando para o objeto solicitado ou sinalizando de modo quase imperceptível, por exemplo) – e aquilo que podemos chamar de *Efeito do Observador*, em que a ordem aparente dos dados brutos não emerge da atividade cognitiva do animal sob

extenso de dados de outros primatas criados em cativeiro, Herbert Terrace e sua equipe (TERRACE *et al.*, 1979) perceberam que aquelas construções que pareciam ser indícios de criatividade linguística poderiam ser explicadas por outros mecanismos, sendo o principal deles a imitação de sinais feitos pelo treinador e que ainda continuavam presentes na memória de trabalho do primata aprendiz.

O mais longo desses filmes, *Teaching Sign Language to the Chimpanzee: Washoe*, mostrava 155 expressões de Washoe, das quais 120 eram expressões de um único sinal. Essas manifestações ocorreram principalmente em sessões de teste de vocabulário. Cada uma das sequências com mais de um sinal de Washoe (24 sequências de dois sinais, 6 de três sinais e 5 de quatro sinais) foram precedidas por uma declaração semelhante ou uma sugestão de sua tutora. Assim, as declarações de Washoe eram adjacentes³³ e imitativas dos enunciados de sua professora. O filme *The first signs of Washoe*, que também mostra Ally (irmão de Nim) [dois chimpanzés] e Koko [uma gorila], revela uma tendência semelhante de o tutor sinalizar antes de o primata produzir os sinais. Noventa e dois por cento dos sinais de Ally e todos os de Koko foram sinalizados pelo tutor imediatamente antes de Ally e Koko sinalizarem (TERRACE *et al.*, 1979: 899)³⁴.

No caso específico ilustrado na Figura 2.8, os autores também se mostram céticos: “*como nós podemos ver na [Fig. 2.8], ambos os sinais de Washoe, time e eat, foram sinalizados por B. Gardner imediatamente antes de Washoe os ter sinalizado. A construção time eat não pode ser considerada uma declaração espontânea por duas razões. Era uma resposta a um pedido de sinalização feito por B. Gardner e continha sinais recém-sinalizados*

análise, mas sim dos princípios cognitivos de organização do próprio observador, que se vê enredado em uma espécie de armadilha kantiana, atribuindo ordem a objetos que não a possuem intrinsecamente. Para uma visão geral sobre o Efeito Clever Hans, cf. Hediger (1981).

³³ Declarações adjacentes são aquelas em que o emissor não respeita a troca de turnos existente em uma conversação normal e acaba interrompendo a fala (ou sinalização) do interlocutor. Padrões imitativos tendem a apresentar declarações adjacentes (simultâneas) ao invés de declarações sucessivas (interpoladas).

³⁴ “The longer of these films, *Teaching Sign Language to the Chimpanzee: Washoe*, showed 155 of Washoe's utterances of which 120 were single-sign utterances. These occurred mainly in vocabulary testing sessions. Each of Washoe's multisign sequences (24 two-sign, 6 three-sign, and 5 four-sign sequences) were preceded by a similar utterance or a prompt from her teacher. Thus, Washoe's utterances were adjacent and imitative of her teacher's utterances. The Nova film, which also shows Ally (Nim's full brother) and Koko, reveals a similar tendency for the teacher to sign before the ape signs. Ninety-two percent of Ally's, and all of Koko's, signs were signed by the teacher immediately before Ally and Koko signed.”

por ela” (TERRACE *et al.*, 1979: 898)³⁵. Uma análise semelhante é proposta para a expressão “*water bird*”, que foi utilizada por Washoe para responder a uma requisição em que ela precisava identificar um cisne – um animal cujo nome não fazia parte de seu banco de sinais.

A resposta de Washoe parece significativa e criativa na medida em que justapõe dois sinais apropriados de uma maneira consistente com a ordem das palavras em inglês. No entanto, não há base para concluir que Washoe estava caracterizando o cisne como um “pássaro que habita a água”. Washoe tinha um longo histórico de situações em que precisava responder à pergunta *what that?* (o que é aquilo?) na presença de objetos como pássaros e trechos d’água. Nesse caso, Washoe pode ter simplesmente respondido à pergunta *what that?*, identificando corretamente um trecho d’água e um pássaro, nessa ordem. Antes de concluir que Washoe estava relacionando o signo *água* (*water*) ao signo *pássaro* (*bird*), deve-se saber se ela regularmente colocava um adjetivo (*water*) antes ou depois de um substantivo (*bird*). Isso não pode ser decidido com base em uma única anedota, não importa o quão convincente essa anedota possa parecer para um observador falante do Inglês (TERRACE *et al.*, 1979: 895-896)³⁶.

Um dos pontos mais significativos do estudo de Terrace foi constatar que o aumento progressivo do comprimento médio das declarações de primatas expostos à ASL não era acompanhado de um aumento correspondente da complexidade sintática e semântica das construções. Assim, havia uma clara assimetria entre a extensão das declarações e o conteúdo informacional veiculado por elas. Apesar de Nim Chimpsky, um dos chimpanzés estudados pela equipe de Terrace, produzir sequências de até 16 sinais justapostos, a quantidade de informação codificada nestas declarações dificilmente ultrapassava aquela veiculada em sequências de dois sinais. Uma amostra das declarações mais frequentes de Nim e uma de suas construções mais extensas encontram-se em (3) e (4), respectivamente.

³⁵ “As can be seen in [Fig. 2.8] both of Washoe's signs (*time*) and (*eat*) were signed by B. Gardner immediately prior to Washoe's having signed them. *Time eat* cannot be considered a spontaneous utterance for two reasons. It was a response to a request to sign by B. Gardner and it contained signs just signed by her.”

³⁶ “Washoe's answer seems meaningful and creative in that it juxtaposes two appropriate signs in a manner consistent with English word order. Nevertheless, there is no basis for concluding that Washoe was characterizing the swan as a “bird that inhabits water”. Washoe had a long history of being asked *what that?* in the presence of objects such as birds and bodies of water. In this instance, Washoe may have simply been answering the question, *what that?*, by identifying correctly a body of water and a bird, in that order. Before concluding that Washoe was relating the sign *water* to the sign *bird*, one must know whether she regularly placed an adjective (*water*) before, or after, a noun (*bird*). That cannot be decided on the basis of a single anecdote, no matter how compelling that anecdote may seem to an English-speaking observer.”

(3) play me
 me Nim
 tickle me
 eat Nim
 more eat
 me eat
 Nim eat
 play me Nim
 eat me Nim
 eat Nim eat
 tickle me Nim
 grape eat Nim
 banana Nim eat
 eat drink eat drink
 eat Nim eat Nim
 banana Nim banana Nim
 drink Nim drink Nim
 banana eat me Nim
 banana me eat banana
 banana me Nim me
 grape eat Nim eat
 Nim eat Nim eat
 play me Nim play

(4) give orange me give eat orange me eat orange give me eat orange give me you

Como podemos perceber em (3) e (4), há um evidente acúmulo de informações repetidas nos enunciados de Nim. A partir de um certo ponto – normalmente, após o segundo ou terceiro sinal –, cada declaração passa a girar em torno dos mesmos elementos iniciais e se estende materialmente de modo cumulativo sem que, no entanto, ocorra um enriquecimento sintático e semântico de sua estrutura. Se considerarmos o conjunto “*give (me) orange*” como o núcleo informacional de (4), notamos que Nim está repetindo esta informação ao longo de toda a sequência, mas de uma forma variada (e um tanto quanto caótica) do ponto de vista sintático – o que mostra o funcionamento de uma mente intensamente ativa, que busca codificar seus desejos, mas que não consegue estruturá-los de modo complexo, uma vez que

(aparentemente) não é dotada de regras que permitam a combinação gradativa e qualitativa – e não apenas quantitativa – dos sinais visuais que formam o seu Léxico. A MLU (*Mean Length Utterance*, Extensão Média do Enunciado) é uma medida de produtividade linguística e se refere ao número médio de morfemas (MLU-m) ou de vocábulos (MLU-w) que um enunciado possui (MARQUES & LIMONGI, 2011). Dado um discurso com (n) enunciados espontâneos, calcula-se a MLU dividindo-se o número total de morfemas (m) ou de palavras (w) por (n): $MLU-m = m/n$; $MLU-w = w/n$.³⁷ Normalmente, os valores da MLU são utilizados como indicadores da competência gramatical de um indivíduo em determinado estágio de seu desenvolvimento linguístico, uma vez que esse incremento material é acompanhado de um incremento gramatical:

A MLU é um excelente índice de desenvolvimento gramatical porque quase todo novo tipo de conhecimento aumenta a extensão da sentença: o número de papéis semânticos expressos, a adição de morfemas obrigatórios, modulações de significado codificadas gramaticalmente, a adição de formas negativas e auxiliares usadas em sentenças interrogativas e negativas e, claro, subordinação e coordenação. Todos têm o efeito comum de incrementar o comprimento da estrutura superficial da sentença (especialmente se a extensão for medida em morfemas, que incluem formas presas – como as desinências – em vez de palavras). (BROWN, 1973: 53-54)³⁸.

Quando Nim produziu o enunciado de 16 sinais representado em (4), sua MLU (o número médio de sinais visuais que ele utilizava por declaração) era de aproximadamente 1,6. Esta imensa disparidade deixa claro, mais uma vez, que o incremento material de seus enunciados não derivava de um incremento de sua competência semântica ou sintática, mas era fruto de mecanismos não gramaticais de extensão, como a iteração enfática (TERRACE *et al.*, 1979: 895). Mas talvez o mais importante a se destacar seja o fato de que a MLU de Nim se manteve relativamente estável ao longo do tempo: se houve algum avanço de seus valores,

³⁷ Para um detalhamento das regras utilizadas na aferição das medidas e no cálculo da MLU, cf. Brown, 1973: 51-58. Para uma visão crítica sobre a correspondência entre MLU e competência gramatical, cf. DeThorne *et al.*, 2005.

³⁸ “The mean length of utterance (MLU) is an excellent simple index of grammatical development because almost every new kind of knowledge increases length: the number of semantic roles expressed in a sentence, the addition of obligatory morphemes, coding modulations of meaning, the addition of negative forms and auxiliaries used in interrogative and negative modalities, and, of course, embedding and coordinating. All alike have the common effect on the surface form of the sentence of increasing length (especially if measured in morphemes, which includes bound forms like inflections rather than words).”

1,6 parece ter sido o limite superior do número médio de sinais encadeados em seus enunciados, o que contrasta violentamente com os valores e com os padrões de variação da MLU relacionados ao desenvolvimento linguístico em seres humanos. Durante as fases iniciais de aquisição linguística, a taxa de variação da MLU (ΔMLU) tende a ser positiva em crianças que apresentam desenvolvimento neurológico normal, indicando um aumento gradativo da competência gramatical ($\Delta MLU > 0 \rightarrow MLU_{\text{final}} - MLU_{\text{inicial}} > 0 \rightarrow MLU_f > MLU_i$). O gráfico da Figura 2.9 mostra as curvas de variação da MLU de cinco indivíduos diferentes: Nim; Eve e Sarah (duas crianças ouvintes-falantes); Ruth e Pola (duas crianças surdas-sinalizadoras). Todas as crianças partem de um patamar similar ao que Nim apresenta aos 26 meses de idade, mas, ao contrário do que ocorre com ele, cujos valores de MLU oscilam sempre em torno de 1,6, todas elas apresentam uma taxa significativa de crescimento da MLU ao longo de seu desenvolvimento.

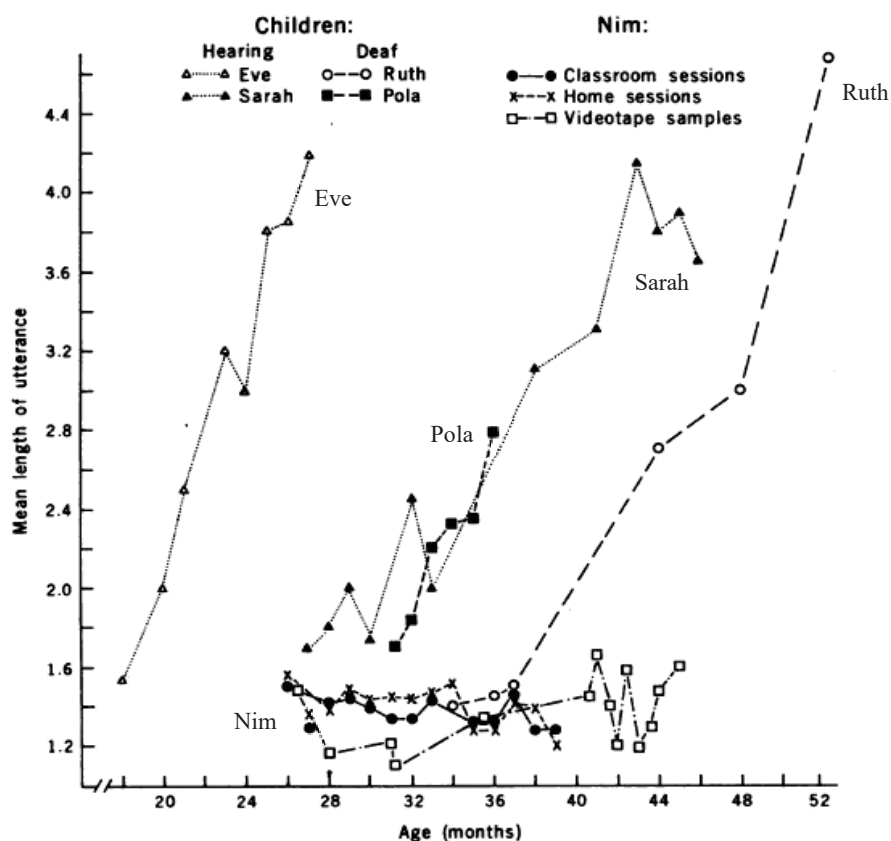


FIGURA 2.9. Gráfico da MLU em função da idade, mostrando a extensão média de enunciados sinalizados do chimpanzé Nim (em três contextos distintos) e de duas crianças surdas e a extensão média de enunciados falados de duas crianças ouvintes. Imagem extraída de Terrace *et al.*, 1979.

A existência de um teto precoce para a MLU de Nim e de outros primatas é mais um dos indícios que apontam para a provável inexistência, em suas mentes/cérebros, de um moto-contínuo gramatical que seja capaz de criar sentenças cada vez mais longas de modo recursivo – uma operação básica que faz parte da competência linguística humana: dada uma sentença (S_1) composta por m morfemas (sendo m um número natural), sempre será possível formar, a partir de S_1 , uma sentença (S_2) constituída por $m + p$ morfemas (sendo p também um número natural), não importa o quão alto seja o valor de m . Assim, mesmo que S_1 possua, por exemplo, 10^3 morfemas, somos virtualmente capazes de construir sentenças cada vez mais extensas aplicando recursos simples, como em (5).

(5) S_2 : *Ana disse que* [S_1].

S_3 : *Maria acha que* [S_2].

S_4 : *João falou que* [S_3].

e assim por diante.

Se sentenças extremamente longas – como o são, teoricamente, as do conjunto (5) – não são produzidas no nosso dia a dia, isto se deve às limitações dos sistemas de desempenho que são acionados durante sua produção e enunciação – como a capacidade da memória de trabalho, a variação do nível de atenção, o condicionamento físico do falante etc. – e não a restrições puramente gramaticais. A MLU baixa e constante de Nim, pelo contrário, parece ser um reflexo dos limites de sua competência gramatical (ou mesmo de sua ausência). Se, entre os seres humanos, não existe algo como “a sentença mais extensa da língua X” (dado que o número de morfemas ou palavras desta sentença sempre pode ser incrementado por mecanismos simples que permitem o espraiamento indefinido de sua extensão), entre os primatas superiores, as construções mais longas de seu sistema de comunicação visual parecem não conseguir ultrapassar o limite de dois ou três sinais combinados.

Em suma, as evidências de que os primatas superiores criam sentenças podem, em cada caso, ser explicadas por referência a processos não linguísticos mais simples. As sequências de sinais produzidas por Nim e por outros primatas podem se assemelhar superficialmente às primeiras sequências multipalavras produzidas por crianças. Mas, a menos que explicações alternativas para as combinações de sinais de um primata sejam eliminadas, em particular o hábito de imitar parcialmente os enunciados recentes dos tutores, não há razão para considerar o enunciado multissinal de um macaco como uma sentença (TERRACE *et al.*, 1979: 900-901).³⁹

³⁹ “In sum, evidence that apes create sentences can, in each case, be explained by reference to simpler nonlinguistic processes. Sequences of signs, produced by Nim and by other apes, may resemble superficially the

RESUMO DO CAPÍTULO E CONSIDERAÇÕES FINAIS

Vimos que, em determinado momento da história evolutiva de nosso planeta, aglomerados bioquímicos adquiriram a capacidade homeostática de manter o equilíbrio interno de seus sistemas e de transferir aos seus descendentes as informações responsáveis por sua construção e seu funcionamento – um salto que fez com que, da matéria inanimada, aflorasse a matéria viva. Vimos também que (i) as informações responsáveis pela arquitetura da vida estão armazenadas nas moléculas de DNA e de RNA, que (ii) todos os organismos do nosso planeta, de todos os tempos, são projeções biológicas de um mesmo código genético universal e que (iii) genes condicionam não apenas a materialidade dos componentes biológicos – como o tecido cerebral de uma pequena abelha –, mas também comportamentos e instintos – como a capacidade desta mesma abelha de (de)codificar informações espaciais e temporais em performances somáticas.

Na sequência, vimos que, dentre todas as formas de vida conhecidas por nós, seres humanos, a nossa espécie parece ser a única capaz de processar linguisticamente e de forma tão complexa o mundo em que vive – seja o mundo da realidade objetiva e imediata (exterior ao indivíduo), seja o mundo subjetivo dos pensamentos e dos sentimentos. De fato, apesar da inegável capacidade cognitiva e comunicativa identificada em outros animais (especialmente naqueles pertencentes ao grupo dos primatas superiores, como chimpanzés e gorilas, mas também em abelhas da espécie *Apis mellifera*, que, como vimos, possuem um sistema comunicacional altamente especializado), e apesar de desconhecermos as reais capacidades linguísticas dos hominídeos que dividiram conosco o planeta em algum momento de nosso passado profundo (como o *Homo neanderthalensis* e o *Homo erectus*), nenhuma das pesquisas desenvolvidas até hoje no campo da aquisição da linguagem conseguiu identificar nos membros de outras espécies a capacidade (ao que tudo indica, exclusivamente humana) de estruturar e significar linguisticamente o mundo em que vivem, codificando-o abstratamente através de um sistema de signos e de regras gramaticais internalizadas. Só nós, seres humanos, podemos, por exemplo, adquirir, logo nos anos iniciais de vida, sem esforço aparente e de modo espontâneo, a língua do ambiente sociointeracional no qual estamos inseridos (desde que não haja nenhuma patologia neurológica severa que comprometa nosso sistema de aquisição); só nós produzimos e compreendemos, a partir de recursos lexicais

first multiword sequences produced by children. But unless alternative explanations of an ape's combinations of signs are eliminated, in particular the habit of partially imitating teachers" recent utterances, there is no reason to regard an ape's multisign utterance as a sentence."

finitos retirados do ambiente linguístico e condicionados em nossa mente, uma infinidade de expressões linguísticas – o que constitui a chamada infinitude discreta (ou criatividade) da competência linguística humana: a nossa capacidade de, através de um conjunto limitado de regras de computação gramatical implícita (ou seja, não consciente), gerar (potencialmente) infinitas combinações entre os elementos (também limitados) que formam nosso léxico mental; só nossa espécie pode atribuir aos signos – aos elementos básicos da comunicação – significados que vão além de seu conteúdo referencial imediato, conferindo polissemia às mensagens transmitidas (um potencial linguístico que, no limite, permitiu à nossa espécie inventar e ficcionalizar).

Gorilas e lobos podem se comunicar com o seu grupo através de vocalizações e expressões corporais (ou mesmo através de sinais bioquímicos, como no caso de formigas e vaga-lumes), mas a mensagem (i) precisa ter sempre a mesma forma geral, sob o risco de não ser compreendida pelos demais membros da comunidade caso ela (a mensagem) se materialize de uma forma não convencional; (ii) não é discreta, ou seja, não pode ser decomposta em componentes mínimos – sua compreensão depende da sua totalidade; e (iii) sempre está associada a um estímulo concreto, exterior ou interior ao indivíduo – sem a percepção de uma ameaça ou sem a percepção da própria fome, nenhum destes animais produzirá o aviso de “perigo” ou sinalizará que está com fome. Nós, humanos, ao contrário, (i) somos capazes de codificar o mesmo conteúdo informacional nas mais diversas formas, sem prejuízo da intercompreensão; (ii) compomos e decompomos as expressões linguísticas em seus constituintes básicos, sendo capazes de inferir e de completar as eventuais lacunas dos enunciados e das mensagens; e (iii) somos capazes de conferir autonomia à linguagem na sua relação com a realidade: inventamos e criamos outras realidades e outros mundos – ou, como disse Nabokov, referindo-se à literatura (mas cuja ideia central se aplica à capacidade de fabulação linguística humana como um todo): *“A literatura não nasceu no dia em que um menino chegou correndo e gritou ‘lobo, lobo’, vindo de um vale neandertal com um grande lobo cinzento em seus calcanhares; a literatura nasceu no dia em que um menino chegou gritando ‘lobo, lobo’, e não havia nenhum lobo atrás dele”* (NABOKOV, 2015: 41). Os seres humanos inventam lobos, transformando-os em linguagem. Um lobo, por sua vez, mesmo que único em sua magnificência, nunca transmitiria à alcateia a ideia de um menino inventado intencionalmente.

CAPÍTULO 3 – INATO VS. ADQUIRIDO

Bebês humanos estão geneticamente programados para adquirir uma língua? Qual é a relação existente entre o sistema cognitivo de base e os dados linguísticos ambientais? Como explicar a assimetria entre um input finito de informações e um sistema potencialmente infinito de geração gramatical? A aquisição da linguagem pode ser explicada apenas por mecanismos gerais da inteligência? Princípios gramaticais estão diretamente escritos em nosso código genético (hipótese inatista forte) ou derivam de um dispositivo neurológico especializado em processar linguagem (hipótese inatista fraca)?

3.1. NATUREZA VS. CULTURA

Se um bebê pirahã nascido às margens do rio Maici fosse adotado por uma família francesa que o levasse com poucos dias de vida para a França, que língua ele falaria após crescer nos subúrbios de Paris? E que língua falaria um bebê francês nascido às margens do Sena se ele fosse adotado ainda com poucos dias de vida por uma família pirahã e crescesse no coração da Amazônia? A não ser que se admita a existência de algum tipo de diferença neurobiológica entre os bebês pirahãs e os franceses que os impeça de adquirir determinados tipos de línguas, as respostas a estas duas perguntas parecem óbvias: independentemente da origem, um bebê imerso na língua francesa falará francês e um bebê imerso na língua pirahã falará pirahã. Esta conclusão decorre de duas premissas básicas: (i) os seres humanos não nascem com uma língua pré-programada em sua mente – ninguém vem ao mundo “sabendo” português, ou “sabendo” francês, ou “sabendo” pirahã (ou “sabendo” qualquer outra língua); (ii) o cérebro de todo e qualquer ser humano encontra-se preparado para adquirir qualquer língua (incluindo as já extintas), uma vez que a faculdade da linguagem – o bioprograma cognitivo (BICKERTON, 1984) que permite esta aquisição – faz parte do patrimônio genético comum à espécie humana.

A aquisição da linguagem, portanto, depende invariavelmente da disposição biológica de nossa espécie e do ambiente linguístico ao qual os seres humanos são expostos durante o período crítico de aquisição – que corresponde, na maioria dos casos, à primeira infância.

Partindo de um estado cognitivo inicial (S_0), o indivíduo em formação, à medida que vai sendo exposto aos dados linguísticos ambientais (D_a) – uma exposição que começa a partir do momento em que o feto desenvolve um sistema perceptual capaz de processar as informações das vozes que até ele chegam –, vai formando sua gramática interna, até o momento em que o sistema linguístico da criança alcança um estado de estabilidade estrutural (S_e), a partir do qual sofrerá pequenas modificações (a maioria delas relacionadas ao *incremento* do Léxico⁴⁰). Em linhas gerais (e abstraindo do processo sua imensa complexidade), a formação de nossa competência linguística – de nossa língua enquanto sistema cognitivo – poderia, portanto, ser equacionada como (1):

$$(1) \quad S_0 + D_a \rightarrow S_e.$$

Nessa equação, tanto os dados linguísticos ambientais (D_a) quanto o sistema inicial de aquisição (S_0) são igualmente indispensáveis à formação do sistema final (S_e). De fato, a ausência de informações ambientais disponibilizadas no momento adequado de maturação neurológica deriva em um sistema gramatical degenerado ($S_0 \rightarrow *S_e$) – como o demonstram os casos de crianças “selvagens” ou que foram intencionalmente isoladas do convívio social (cf. LUST, 2006: 93-96; GLEITMAN & NEWPORT, 1995). Por outro lado, sem um sistema cognitivo de base que processe os dados e oriente a aquisição, a formação de S_e se torna impossível ($D_a \rightarrow *S_e$) – o que explica por que outras espécies, mesmo quando submetidas a um processo de ensino sistemático da língua, com exposição ostensiva aos dados linguísticos, nunca conseguiram sequer se aproximar da competência linguística (S_e) que uma criança alcança ao final do processo de aquisição: sem um sistema S_0 especialmente dedicado à “absorção” e ao processamento do *input* linguístico, os dados ambientais são pouco mais do que vestígios sonoros ou visuais que batem e se despedaçam cegamente no tecido cerebral.

Dentro desse quadro teórico (cujas bases podem ser encontradas em Chomsky, 1975: Cap. 1), o desequilíbrio entre os dados linguísticos retirados do ambiente (D_a) e a capacidade de produção que o sistema linguístico apresenta ao final do processo de aquisição (S_e) evidencia o chamado *problema lógico da aquisição da linguagem*: como um *input* finito

⁴⁰ A *aquisição* do Léxico, por sua vez, começa aproximadamente aos seis meses de idade e envolve duas tarefas não triviais: (i) “recortar e extrair” do fluxo contínuo da fala aqueles blocos de sons que servirão de base para delimitar as unidades do Léxico (vocábulos ou morfemas); (ii) associar a cada uma dessas formas fonológicas um significado (ou um referente, segundo a hipótese *word-to-world mapping procedure*). Para mais detalhes sobre a natureza e o grau de complexidade dessas tarefas lexicais, que, provavelmente, também são determinadas geneticamente, cf. GUASTI, 2002, cap. 3.

(formado por estímulos ambientais limitados) pode dar origem a um *output* infinito (o comportamento linguístico em sua plenitude)? Esse problema, no limite, dialoga com outras duas assimetrias. Por um lado, durante a aquisição do conhecimento linguístico, há uma *riqueza* (ou *excesso*) de *estímulos ambientais*: na infância (talvez, já na vida intrauterina (FIFER & MOON, 1989)), o experienciador é colocado diante de uma profusão de estímulos sensoriais, linguísticos e não linguísticos, mesclados de forma caótica e constituindo um emaranhado de informações de diferentes naturezas. Mesmo assim, ao final do processo de aquisição, a criança domina plenamente todos os componentes da gramática de sua língua. Como ela é capaz de distinguir quais informações sensoriais são pertinentes e quais são descartáveis com relação à formatação de sua gramática? Há uma clara assimetria entre a *quantidade bruta* de informações ambientais e a *quantidade (linguisticamente) útil* de informações ambientais⁴¹. Por outro lado, existe uma relação assimétrica entre o sistema

⁴¹ O mapeamento dos sinais acústicos e a seleção das informações linguisticamente relevantes derivam de um sistema especializado que pode se relacionar com o sistema geral de processamento sonoro de duas formas distintas: “[podemos] assumir duas relações possíveis entre um sistema especializado de percepção da fala e um sistema generalizado de reconhecimento de objetos auditivos [...]. Os sistemas podem ser construídos um sobre o outro de forma que a percepção da fala seja um caso especial de reconhecimento auditivo. A fala, nesta visão, é diferente do processamento auditivo generalizado porque lida com sinais familiares, frequentes, ‘superaprendidos’. As previsões dessa visão incluem (i) que todos os sinais são inicialmente processados pelos mecanismos gerais antes de serem elaborados como traços específicos da fala e (ii) que danos aos mecanismos gerais de reconhecimento de padrões necessariamente levam a uma percepção de fala comprometida. Alternativamente, a percepção da fala e o reconhecimento de objetos auditivos não linguísticos podem funcionar em paralelo. Nesta visão, os sistemas dependem de mecanismos compartilhados no início do fluxo de processamento, mas fazem uso de arquiteturas especializadas na construção das percepções relevantes. Além disso, eles podem ser danificados seletivamente. [...] Ambas as propostas são consistentes com a ideia de que os principais objetivos computacionais do sistema auditivo são a localização de sons, o reconhecimento de vocalizações coespecíficas (como aquelas relativas à fala) e o reconhecimento de outros padrões auditivos que ocorrem no ambiente acústico” (POEPPEL, 2001: 680). (Texto original: “We assume two possible relations between a specialized speech perception system and a generalized auditory object recognition system [...]. The systems could build on one another such that speech perception is a special case of auditory recognition. Speech, on this view, is distinct from generalized auditory processing because it deals with familiar, frequent, ‘overlearned’ signals. Predictions of this view include (i) that all signals are initially processed by the generalized mechanisms before being elaborated as speech percepts and (ii) that damage to general pattern recognition mechanisms necessarily leads to compromised speech perception. Alternatively, speech perception and nonspeech recognition may function in parallel. On this view, the systems rely on shared mechanisms early in the processing stream but make use of specialized architecture in the construction of the relevant percepts. Moreover, they can be damaged selectively. [...] Both proposals are consistent with the idea that the main

linguístico altamente organizado e sofisticado alcançado pela criança ao final do processo de aquisição e os dados linguísticos do ambiente, que, apesar de abundantes quantitativamente, são pobres qualitativamente, na medida em que (i) não mostram de forma explícita as regras que subjazem às suas estruturas (a criança tem acesso consciente aos enunciados materiais de uma língua, mas não às estruturas e regras de formação desses enunciados) e (ii) não contêm as *informações negativas* do sistema, i.e., não contemplam aquelas sentenças que nunca são produzidas, mas que, mesmo assim, as crianças *sabem* que são impossíveis em sua língua⁴². A título de ilustração, pensemos por um instante em como nós, desde muito cedo, conseguimos mapear adequadamente as relações de correferenciação entre DPs em construções como (2) e (3):

(2) [O vovô]^B falou que [ele]^B pegou a bola.

(3) [Ele]^B falou que [o vovô]^{*B} pegou a bola.

Em (2), o pronome *ele* se encontra coindexado à expressão-R *o vovô*, indicando que ambos os sintagmas possuem a mesma referência/conteúdo referencial. Já em (3), a coindexação (artificial) entre o pronome *ele* e a expressão-R *o vovô* resulta em uma construção agramatical que não é licenciada pelo sistema (de fato, a única leitura possível de

computational goals for the auditory system are the localization of sounds, the recognition of conspecific vocalizations such as speech, and recognition of other auditory patterns occurring in the acoustic environment.”).

⁴² Uma descrição mais completa da assimetria que marca a relação entre o conhecimento linguístico alcançado pela criança ao final da fase de aquisição e as informações disponíveis nos estímulos linguísticos ambientais pode ser encontrada resumidamente em Burraco (2012: 37). A aquisição da língua materna é marcada por “sua significativa (i) velocidade (a aquisição ocorre a uma velocidade anormalmente alta), (ii) confiabilidade (a aquisição acaba ocorrendo efetivamente em todos os casos, exceto nos patológicos), (iii) produtividade (a capacidade linguística adquirida permite gerar e compreender um número ilimitado de sentenças), (iv) seletividade (a gramática finalmente adquirida é, *a priori*, apenas mais uma entre muitas outras alternativas, que, no entanto, resultariam em erro), (v) indeterminação (a teoria gramatical finalmente desenvolvida nasce de um conjunto de dados intrinsecamente indeterminados), (vi) convergência (as gramáticas adquiridas por todos os membros de uma dada comunidade de fala são semelhantes) e (vii) universalidade (os sistemas gramaticais adquiridos em qualquer comunidade de fala compartilham semelhanças substanciais). Mas, da mesma forma, o argumento da “pobreza de estímulo” também se apoia em certas características do próprio estímulo linguístico que elicita essa aquisição, em particular, seu caráter (i) finito, (ii) idiossincrático (isto é, diverso), (iii) incompleto (no sentido de que há inúmeras construções que a criança nunca ouve) e (iv) degenerado (isto é, que contém vários tipos de erros), bem como o fato de não incluir (v) qualquer tipo de recompensa (o que excluiria a ideia de que a aquisição da linguagem consiste em um processo de aprendizagem condicionado), nem (vi) qualquer tipo de especificações negativas (ou seja, de indicações sobre quais construções são agramaticais).”

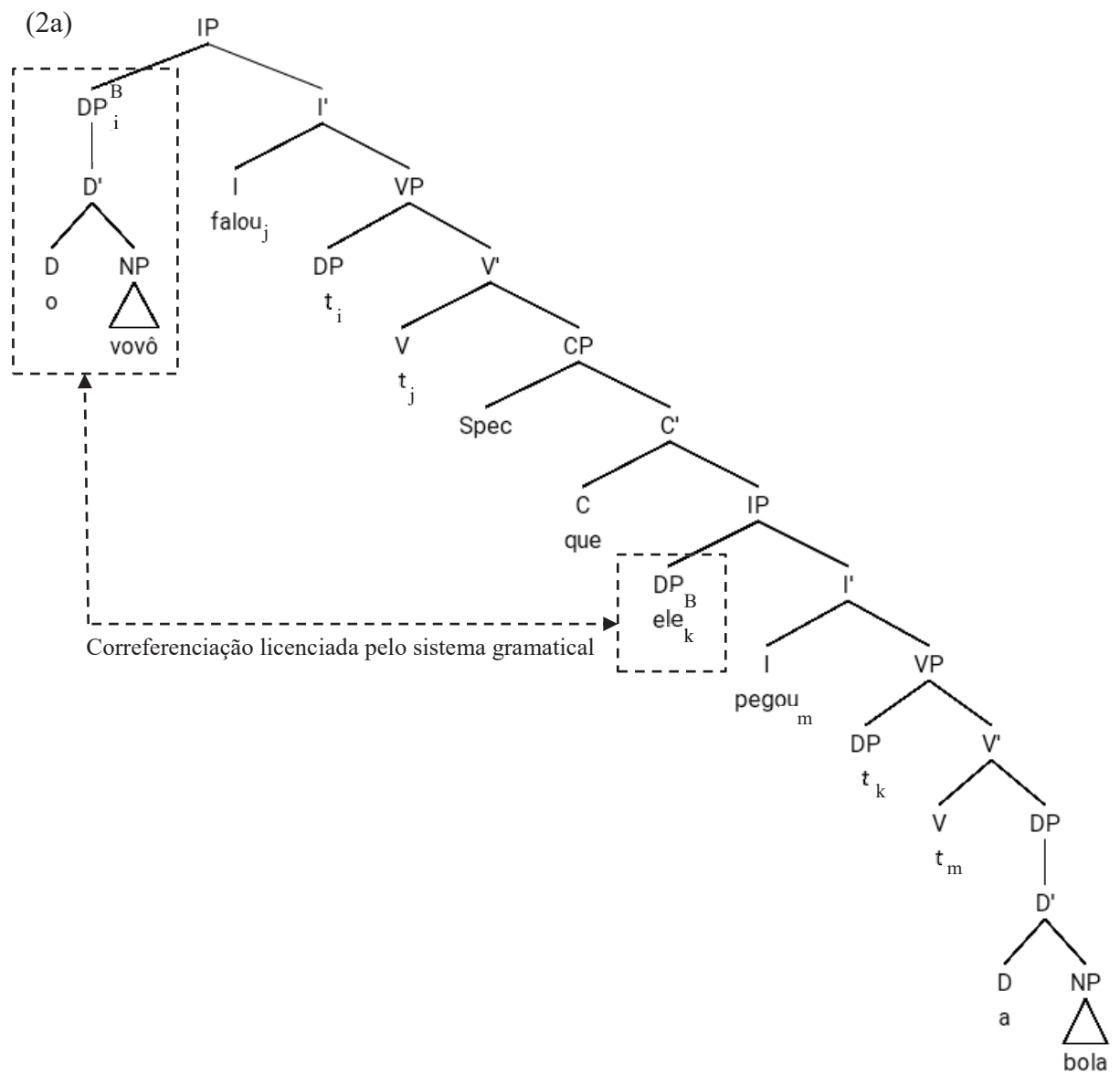
uma construção como (3) é aquela em que a expressão-R e o pronome são interpretados disjuntivamente: [Ele]^A falou que [o vovô]^B pegou a bola). Este contraste de gramaticalidade entre as construções indica que os correferentes não se distribuem de modo aleatório ao longo das sentenças e que há um conjunto de regras e restrições que organizam sua distribuição e ligação. Mas será que esse conjunto de regras de correferenciação é aprendido ou deriva de um bioprograma já “instalado” em nosso sistema cognitivo? No limite, estamos discutindo um problema que faz parte de uma oposição maior: o que, no Humano, vem da Natureza vs. o que, no Humano, vem da Cultura. Em que pese não ser um problema simples, há pelo menos um argumento forte a favor da hipótese inatista do sistema de correferenciação.

Uma vez que (i) as crianças só têm acesso às evidências positivas dos dados linguísticos efetivamente produzidos (que incluem sentenças estruturalmente semelhantes a [2]) e que (ii), mesmo sem jamais terem escutado as formas negativas (ou seja, *não produzidas*) de correferenciação, elas sabem que construções como (3) são impossíveis na língua, é preciso postular a existência de um sistema inato que guie a criança durante o processamento e o mapeamento dos correferentes, indicando a ela quais são as construções gramaticais/licenciadas ou agramaticais/bloqueadas de sua língua. Do contrário, sem esse sistema inato de restrições, como poderia a criança saber que a estrutura hipotética (3) – uma sequência cuja configuração de correferentes nunca está disponível nos dados linguísticos a que o *infans* tem acesso – é uma construção impossível em sua língua, apesar de ela possuir, do ponto de vista estritamente semântico, um conteúdo proposicional (artificialmente construído através da coindexação) que é exatamente o mesmo da sentença (2)?

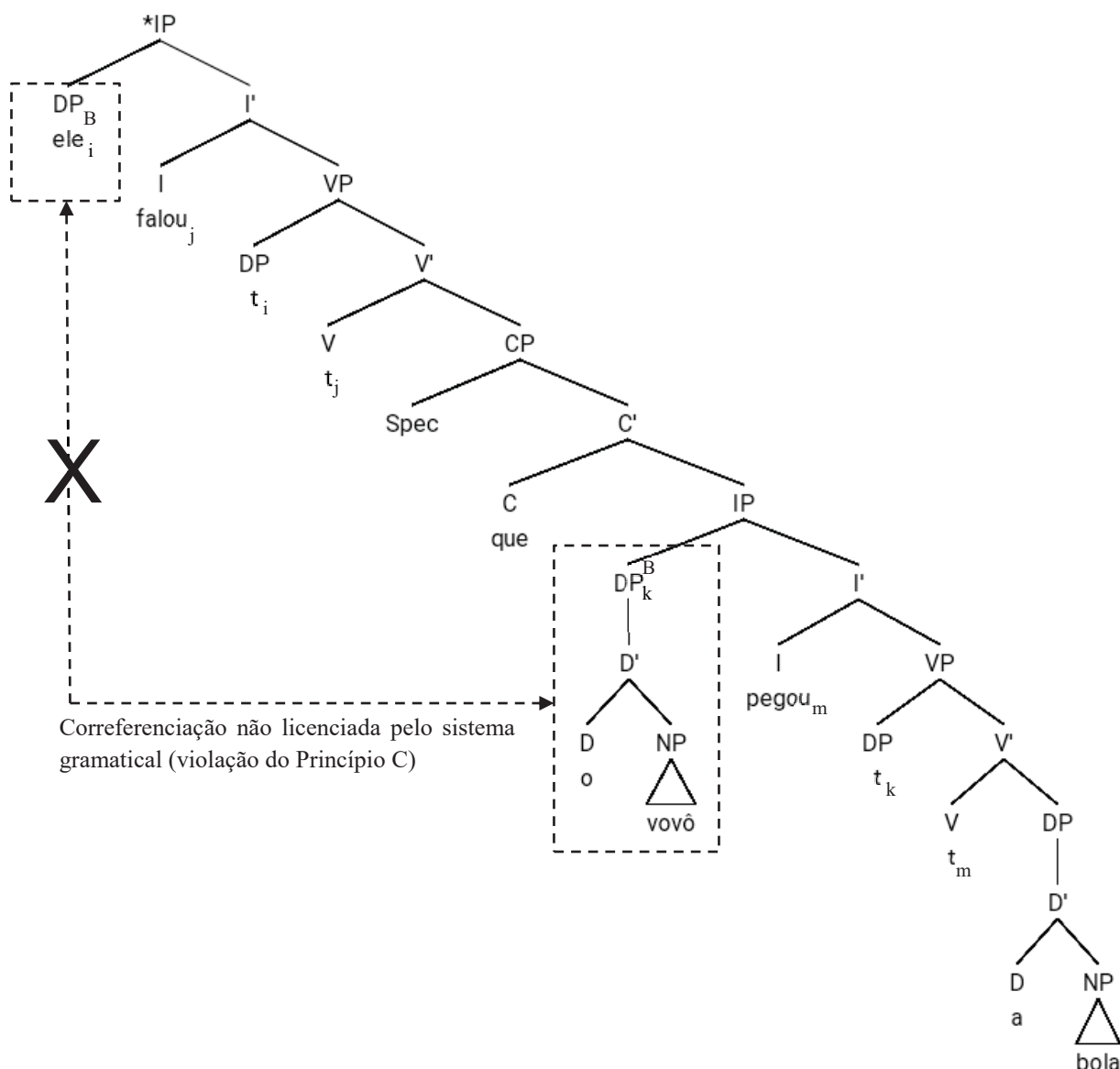
No capítulo 6, quando realizarmos um balanço final do problema da relação entre genes e linguagem, retomaremos alguns pontos do debate sob a perspectiva do *Programa Minimalista* (PM), a versão mais atual da Gramática Gerativo-Transformacional (GGT). No entanto, para os fins da presente discussão, lançaremos mão de construtos teóricos basilares – como Princípios e Parâmetros – da forma como eles foram propostos em um momento anterior da teoria – mais especificamente, dentro do quadro teórico da *Government and Binding Theory* (GB ou TRL: *Teoria da regência e ligação*). O motivo para essa escolha está na natureza mais “audaciosa” (e arriscada) dessa fase da GGT, que apostava em um conjunto mais ou menos amplo de diretrizes gramaticais inscritas de alguma forma no genoma dos seres humanos, o que nos permite ilustrar determinados pontos críticos do problema com mais clareza do que se adotássemos uma argumentação orientada pelo Minimalismo. De certa forma, a GB nos habilita a fazer afirmações mais “categóricas” e explícitas sobre

determinados princípios gramaticais inatos (*hipótese inatista forte*); o PM, por sua vez, busca otimizar o corpo teórico do programa gerativista e aposta em diretrizes gramaticais que *derivam* de princípios cognitivos mínimos codificados geneticamente – como a capacidade de recursão –, o que corresponde a uma teorização mais comedida e ajustada do ponto de vista neurobiológico (*hipótese inatista fraca*). Ainda neste capítulo, discutiremos um pouco mais sobre a natureza epistemológica dessas duas hipóteses. Feita essa consideração, voltemos à questão do mapeamento “instintivo” de correferentes que as crianças executam ao processar os DPs de uma sentença.

No quadro da GB, tal sistema de correferenciação, por emergir de mecanismos inatos, foi pensado como um constituinte gramatical presente na base de todas as línguas naturais, sendo composto por três princípios universais que regulam a distribuição de anáforas (Princípio A), pronomes (Princípio B) e expressões referenciais (Princípio C) dentro das sentenças (CHOMSKY, 1981). Apesar de os exemplos (2) e (3) exigirem a computação de pronomes (Princípio B), mostrando que atividade gramatical efetuada pela criança é ainda mais complexa, para os fins aqui propostos, vamos nos ater ao Princípio C, que explica a gramaticalidade de (2) e a agramaticalidade de (3). As estruturas-S (superficiais) de (2) e (3) estão representadas segundo o modelo da GB em (2a) e (3a), respectivamente (os índices subscritos [em letras minúsculas] indicam os movimentos e os sobrescritos [em letras maiúsculas], as coindexações. As representações arbóreas seguem as diretrizes de construção da Teoria X-barra):



(3a)



O Princípio C estabelece que expressões referenciais (expressões-R) têm autonomia referencial e precisam estar livres na sentença – ou seja, expressões-R não podem ser c-comandadas⁴³ por elementos que possuam o mesmo índice referencial. Em (2a), a expressão-R *o vovô* ocupa a posição de especificador do IP da sentença matriz e não é c-comandada por nenhum elemento correferencial, obedecendo ao Princípio C. Em (3a), entretanto, o DP *o vovô* ocupa a posição de sujeito da sentença encaixada e é c-comandado por um elemento coindexado, o pronome *ele*, o que viola o Princípio C e, portanto, torna a construção agramatical. Apesar da complexidade e da ausência de evidências negativas, crianças, desde

⁴³ X c-comanda Y se o nó que domina imediatamente X também domina Y.

muito cedo, são capazes de mapear adequadamente a distribuição de anáforas, pronomes e expressões-R ao longo das sentenças e sabem instintivamente que relações de correferenciação como a que foi sugerida em (3) não são licenciadas pelo sistema gramatical da língua. Em (3), a exigência de autonomia que o Princípio C impõe à expressão-R *o vovô* faz com que esse DP e o pronome que o c-comanda (*ele*) sejam processados disjuntivamente pela competência gramatical da criança, gerando a interpretação compulsória (4)⁴⁴.

(4) [Ele]^A falou que [o vovô]^B pegou a bola.

Como solucionar o problema lógico da aquisição da linguagem (*input* finito → *output* infinito) e equacionar (i) a assimetria existente entre a quantidade bruta de informações ambientais captadas pelos sistemas de percepção da criança e a quantidade útil de informações linguísticas (problema do excesso de estímulo) e (ii) a assimetria qualitativa existente entre os dados linguísticos ambientais e a competência alcançada pela criança ao final do processo de aquisição (Problema de Platão⁴⁵)? Seria possível postular a existência de um conjunto de “conhecimentos inatos” que seriam capazes de filtrar, organizar e integrar os dados incompletos da experiência? No âmbito da GGT, formulou-se a *hipótese do inatismo linguístico* para dar conta dessas assimetrias.

De acordo com essa hipótese, somente o ser humano é capaz de filtrar os dados da sua experiência de modo a construir uma competência linguística porque essa é uma dotação genética da nossa espécie. O inatismo linguístico sustenta, portanto, a tese de que a predisposição para a linguagem é uma característica biológica do *Homo sapiens* (KENEDY, 2013: 74).

Desta forma, segundo a GGT, a criança traz consigo um sistema inato – denominado *faculdade da linguagem* (S₀) – que é capaz de analisar os dados (D_a) da língua-ambiente (também chamada de língua-E: um conjunto de (e)xtensões (e)xplícitas (e)xternas ao

⁴⁴ Para uma visão geral sobre a teoria da regência e ligação, cf. HAEGEMAN (1994); LOBATO (1986). Para uma revisão crítica sobre os testes experimentais utilizados em estudos sobre aquisição do Princípio C, cf. LUST *et al.* (1992). Para um breve detalhamento do processo de aquisição do Princípio C no português brasileiro, envolvendo questões metodológicas, cf. GROLLA (2013).

⁴⁵ Convencionou-se chamar a assimetria existente entre o conhecimento complexo e organizado que possuímos e a experiência limitada bruta a que estamos condicionados de Problema de Platão. Em *Knowledge of language: Its nature, origin and use* (1986), Chomsky reformula o problema em termos de aquisição da linguagem.

indivíduo) e de moldar cognitivamente uma competência linguística (S_e) (também chamada de língua-I: um sistema (i)ntensional (i)mplicito (i)nterno ao (i)ndivíduo).

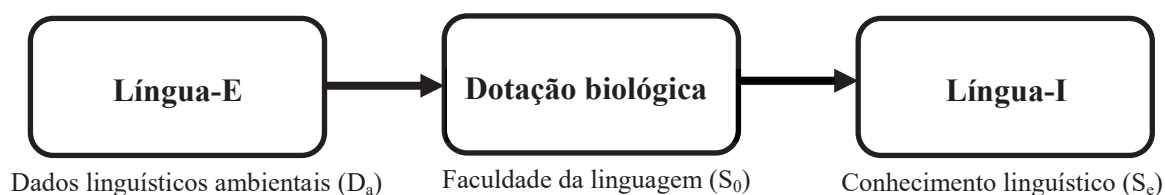


FIGURA 3.1. A hipótese inatista sobre a faculdade da linguagem (adaptado de KENEDY, 2013: 74).

Segundo a hipótese inatista, portanto, os seres humanos estão pré-programados geneticamente para poder adquirir uma língua. Mas qual seria a real natureza da faculdade da linguagem?

No quadro da Teoria de Princípios e Parâmetros (CHOMSKY, 1981; 1982), a faculdade da linguagem foi teorizada como um conjunto de especificações denominado *Gramática Universal* (GU). Os *princípios* são as constantes linguísticas universais (comuns a todas as línguas e presentes na estrutura cognitiva de todos os seres humanos); já os *parâmetros* são as variáveis que, através das experiências linguísticas de cada indivíduo em sua comunidade de fala, assumem valores específicos. A fim de ilustrar a distinção entre estes dois tipos de diretrizes, vamos analisar como o Princípio de Projeção Estendido e o chamado Parâmetro do Sujeito Nulo⁴⁶ se realizam em uma sentença como (5). Para isso, vamos assumir uma arquitetura da linguagem ancorada em um Léxico mais robusto (*hipótese lexicalista*).

(5) [IP₁ Maria disse que [IP₂ comprou um avião]].

Muito mais do que simples pares de som e significado, itens lexicais verbais codificam em sua estrutura argumental uma série de propriedades de seleção e de atribuição de funções temáticas. Assim, um predador verbal como *comprar* se encontra programado para selecionar como argumento interno um DP que possa receber o papel temático (papel- θ) de TEMA e como argumento externo um DP que esteja apto a receber o papel- θ de AGENTE.

⁴⁶ A possibilidade de existência de um sujeito nulo (i.e., sem matriz fonêmica) é uma das propriedades do Parâmetro *pro-drop* (do inglês, *pronoun dropping*: supressão do pronome). Para mais detalhes, cf. CHOMSKY, 1981:240-248.

(6) Grade argumental de *comprar*:

categoria	[–Nome, +Verbo]
número de argumentos	[____ comprar ____]
c-seleção (seleção categorial)	[DP, DP]
s-seleção (seleção semântica)	[AGENTE, TEMA]

Segundo o *Princípio de Projeção* (PP), “as informações lexicais (tais como os papéis temáticos) são sintaticamente representadas em todos os níveis” (CARNIE, 2006: 225)⁴⁷. Assim, ao serem projetados na sentença, os itens lexicais trazem consigo as exigências de suas respectivas grades argumentais e as requisições impostas por elas. Qualquer violação relacionada à distribuição de papéis temáticos, por exemplo, resulta em agramaticalidade (cf. a construção **Maria comprou* [_], em que a agramaticalidade decorre do fato de *comprar* ter dois papéis temáticos para descarregar (AGENTE e TEMA), mas apenas um argumento (*Maria*) para recebê-los na representação sintática⁴⁸). Já o *Princípio de Projeção Estendido* (EPP, do inglês *Extended Projection Principle*) estabelece que “todas as orações têm sujeito, i.e., o especificador do IP precisa ser preenchido sempre por um DP ou por um CP” (adaptado de CARNIE, 2006: 227)⁴⁹. Logo, uma construção como *[_] *comprou um avião* é considerada agramatical porque não apresenta em sua estrutura um argumento que possa receber o papel- θ de AGENTE (o que constitui uma violação ao PP) e que possa exercer a função de sujeito da oração (o que constitui uma violação ao EPP). Note, entretanto, que a gramaticalidade da sentença (5) indica que ambos os princípios estão sendo satisfeitos na oração encaixada (IP₂), o que nos permite postular a existência de um tipo de argumento nulo, chamado *prozinho* (*pro*), que, apesar de não possuir matriz fonêmica, é processado pelo sistema de interpretação da sentença como se fosse um representante da classe dos pronomes

⁴⁷ “*The Projection Principle*: Lexical information (such as theta roles) is syntactically represented at all levels.”

⁴⁸ Obviamente, a construção *Maria comprou* é gramatical nos casos em que o argumento interno do verbo *comprar* é uma *ec* (*empty category*: categoria vazia) cujo conteúdo referencial pode ser recuperado através do mapeamento do contexto.

Interlocutor A: *E o avião?*

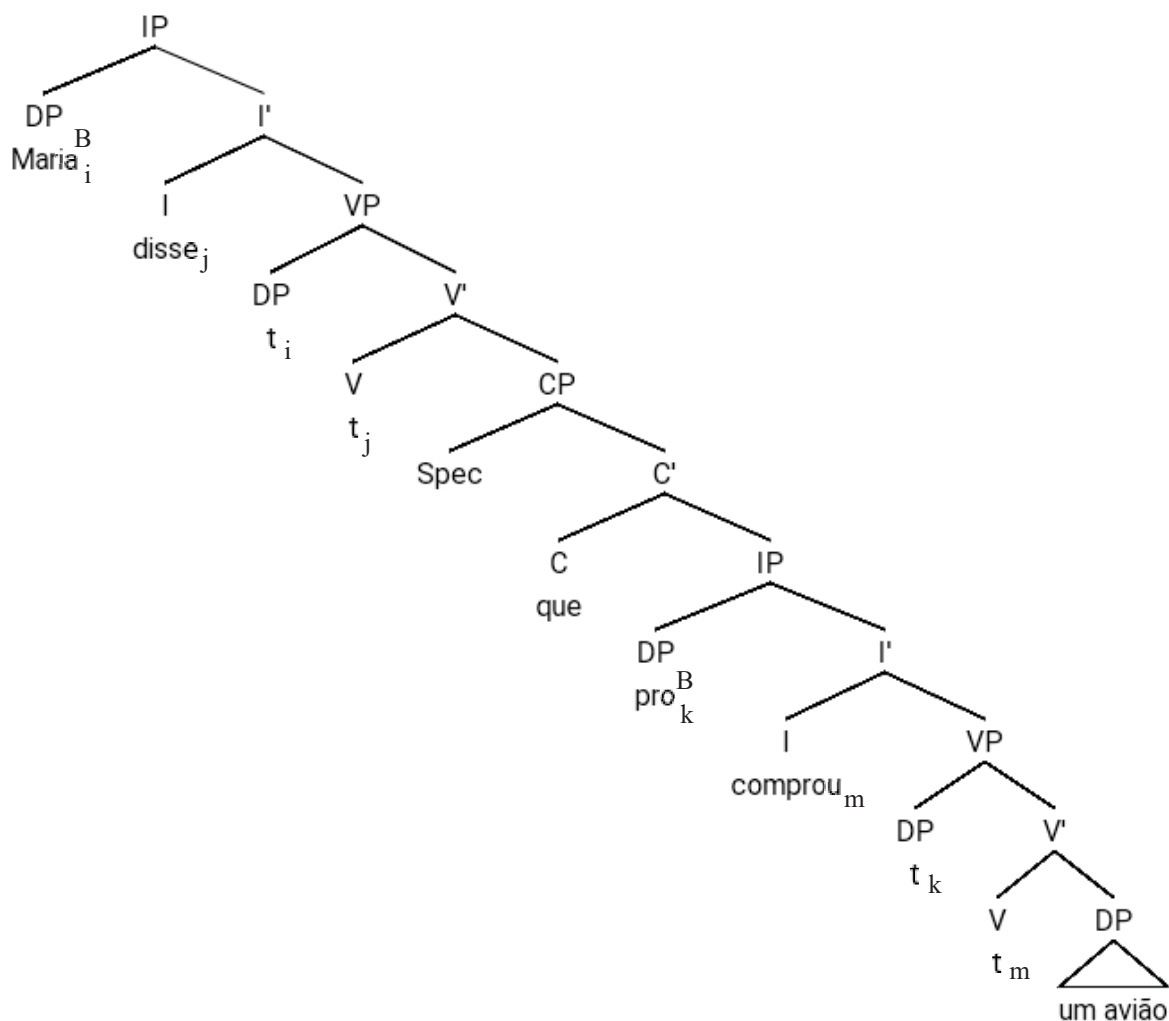
TÓPICO DISCURSIVO (TOP): AVIÃO

Interlocutor B: *A Maria comprou.* → [TOP^K [IP *A Maria comprou ec*^K]].

⁴⁹ “*Extended Projection Principle* (EPP): All clauses must have subjects. (i.e. the specifier of TP must be filled by a DP or CP).”

coindexado ao DP *Maria* (cf. *Maria_i disse que pro_i comprou um avião* vs. *Maria_i disse que ela_i comprou um avião*). Em (5), o *pro* recebe o papel- θ de AGENTE na estrutura-D (estrutura profunda) e se movimenta em direção à posição de especificador de IP₂, a fim de receber Caso Nominativo e exercer a função de sujeito da oração encaixada (cf. o diagrama arbóreo 5a).

(5a)



O EPP, portanto, faz parte do componente computacional da linguagem humana e é responsável por impor à estrutura que está sendo derivada uma exigência quanto ao preenchimento do Spec de IP, gerando construções linguísticas que possuam sujeito, seja ele lexicalmente realizado ou não. A possibilidade de existência de um sujeito sem matriz fonêmica, no entanto, é determinada parametricamente. Línguas não *pro-drop*, como o Francês e o Inglês, exigem a presença de um sujeito lexicalmente explícito (cf. 7 e 8); línguas parcialmente *pro-drop*, como o PB, admitem a existência de pronomes nulos em

determinados contextos, como em (5)⁵⁰; línguas plenamente *pro-drop*, como o Italiano, em que a realização plena do sujeito se torna redundante, uma vez que as desinências verbais já se encarregam de diferenciar as pessoas do discurso (RIZZI, 1988), admitem a presença de sujeitos nulos realizando todas as pessoas discursivas e em configurações sintáticas variadas (cf. 9 e 10).

- (7) [Maria]^B said [she]^B bought a plane.
- (8) [Maria]^B a dit qu'[elle]^B avait acheté un avion.
- (9) [Maria]^B ha detto che [pro]^B comprerà un aereo. [Maria disse que comprará um avião].
- (10) a. *pro* verrà.
 b. **pro* will come.
 c. **pro* virá. (ele/ela virá).

(Adaptado de RIZZI, 1988: 71)

O entrelaçamento entre princípios (como o de projeção estendido) e parâmetros (como o do sujeito nulo) permite conciliar em um mesmo modelo teórico os denominadores comuns que atravessam todas as línguas e a variabilidade de gramáticas que derivam da GU. Da mesma forma que a variação fenotípica de uma determinada característica biológica ocorre

⁵⁰ As mudanças pelas quais vem passando o quadro pronominal do PB – que envolvem (i) a substituição quase que total do pronome *tu* por *você*; (ii) a eliminação do pronome *vós* e sua substituição por *vocês* e (iii) a ascensão da forma *a gente*, em detrimento do pronome *nós* – têm provocado uma redução do paradigma verbal e uma proliferação de “formas sincréticas”: (eu) *compro*; (você) *compra*; (ele/ela) *compra*; (a gente) *compra*; (vocês) *compram*; (eles/elas) *compram*. Segundo Duarte (1995), o empobrecimento do paradigma verbal do PB tem feito com que a concordância deixe de exercer a função de diferenciação e identificação das pessoas do discurso, mostrando-se insuficiente para a recuperação do sujeito. Como consequência dessa redução do quadro flexional, os sujeitos pronominais lexicalmente realizados vêm aos poucos ocupando os espaços que eram antes preenchidos pelos sujeitos pronominais nulos (*pro*). Aparentemente, quanto menos diversificado é o quadro de flexões verbais que realizam a concordância, maior é a necessidade de sujeitos plenos (e menor é a funcionalidade de sujeitos nulos). Desta forma, a existência de valores flutuantes para o parâmetro do sujeito nulo tem feito com que o PB atual se configure como um sistema gramatical misto (ou, na tipologia proposta por Holmberg & Roberts (2009), como uma língua *pro-drop* parcial) que talvez esteja em uma fase de transição rumo a uma língua não *pro-drop*. Apesar do seu quase desaparecimento quando olhamos para o quadro flexional do PB atual, o sujeito nulo continua ativo em determinados contextos, como em orações encaixadas (*A Maria disse que [pro] vai comprar um submarino quando sair o pagamento*) e em construções que giram em torno de um tópico discursivo (*Parece que as pessoas não aprendem com os próprios erros. Só isso explica porque [pro] continuam votando nos mesmos candidatos*).

dentro de certos limites condicionados geneticamente⁵¹, o espaço em que uma língua pode variar é limitado aprioristicamente pelas restrições da GU. Como todo componente biológico da natureza, uma língua só pode variar a partir de uma base comum pré-programada geneticamente (i.e., uma língua natural só pode ser aquilo que o código genético humano e a biologia de nosso cérebro permitem. Os frutos que uma macieira gera a partir de água, nutrientes e luz podem variar em cores, tamanhos e sabores, mas serão sempre maçãs; os produtos que o cérebro humano gera a partir de dados linguísticos ambientais podem variar circunstancialmente, mas serão sempre línguas naturais). A GU, portanto, configura-se como o estado inicial (S_0) do processo de aquisição e é a matriz responsável pela internalização de qualquer gramática humana, fornecendo os princípios com os quais a criança é capaz de construir, através do processamento dos dados linguísticos do ambiente e da fixação dos parâmetros, sua gramática interna (S_e).

Assumir a existência de um estágio inicial não nulo, determinado pela GU, significa assumir que a criança não inicia seu processo de aquisição da língua como *tabula rasa* na qual as informações linguísticas serão inscritas passivamente⁵². Durante o processamento de uma

⁵¹ Na natureza, os casos de *variação fenotípica descontínua*, em que os fenótipos são radicalmente distintos um do outro (como, por exemplo, *albino* vs. *não albino*), são pouco frequentes. O mais comum é que um mesmo genótipo condicione um *continuum* de fenótipos, dentro de certos limites. O espectro de variação fenotípica que uma determinada configuração gênica pode expressar é chamado de *norma de reação gênica*.

⁵² Talvez uma das passagens mais conhecidas sobre a concepção de mente associada à ideia de *tabula rasa* seja a que abre a seção 2 (“*All ideas come from sensation or reflection*”) do primeiro capítulo do Livro II de *An Essay Concerning Human Understanding*, de John Locke: “Suponhamos então que a mente seja, como dissemos, um papel em branco, desprovido de todos os caracteres e sem ideias: – Como ela é preenchida? De onde vem aquele vasto estoque que a fantasia ocupada e ilimitada do homem pintou nela com uma variedade quase infinita? De onde vêm todos os materiais da razão e do conhecimento? A isso respondo em uma só palavra: da experiência. Nela todo o nosso conhecimento está fundado; e dela todo nosso conhecimento deriva em última instância.” (LOCKE, 1999 [1690]: 87. Cf. o texto original ao final desta nota). Para Locke, todo o conhecimento – enquanto realidade mental – advém da experiência do indivíduo e todas as ideias derivam (i) das impressões captadas dos objetos externos (sensações) e (ii) das operações internas de processamento mental (reflexões). Esse empirismo radical, que nega a existência de princípios inatos de qualquer natureza, dialoga profundamente com a filosofia política do autor e com sua oposição às verdades eternas e aos direitos inatos nos quais se fundamentavam os poderes das realezas absolutistas de seu tempo. John Locke, assim como David Hume e Francis Bacon, pertence a uma tradição filosófica cujo princípio condutor, baseado na obra de Tomás de Aquino, encontra-se sintetizado no aforismo “*nihil est in intellectu quod non fuerit prius in sensu*” (“Não há nada no intelecto que não tenha passado antes pelos sentidos”). Um pequeno acréscimo proposto por Gottfried Wilhelm Leibniz – “*nisi*

língua qualquer, a faculdade da linguagem, entendida como um dos módulos do sistema cognitivo humano, fornece à criança, logo de saída, um conjunto de princípios (universais) e de variáveis (que serão formatadas pelos dados do ambiente) com os quais ela se orienta “instintivamente” (ou seja, de modo não-consciente). Neste modelo, os dados linguísticos ambientais funcionam como gatilho (*trigger*) do processo de aquisição, sendo analisados por um sistema cognitivo inato especializado (GU) – o que permite uma analogia com o mecanismo geral de expressão fenotípica, em que o *trigger* são os fatores ambientais (LIGHTFOOT, 1989: 321):

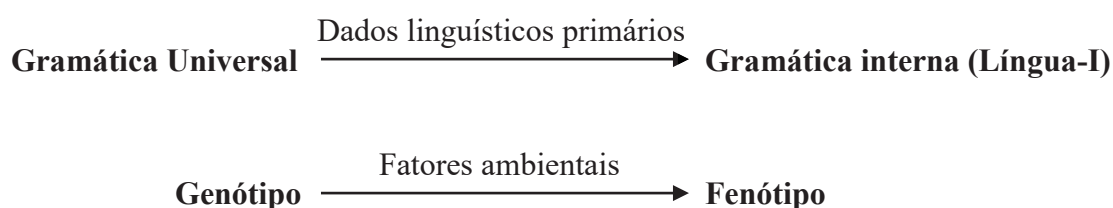


FIGURA 3.2. Mecanismos de expressão fenotípica. A gramática interna (fenótipo) de uma criança é derivada de uma GU geneticamente determinada, alimentada por dados linguísticos.

Esta aproximação entre o mecanismo gerativo de aquisição da linguagem e o mecanismo de expressão dos genes, apesar de ilustrativa, precisa ser analisada dentro de certos limites. De fato, a transformação da GU em uma gramática interna se configura como um processo cognitivo concebido dentro do quadro teórico do gerativismo. A expressão gênica, por sua vez, configura-se como um processo bioquímico de regulação da atividade biológica através da produção de proteínas – o que inclui o controle da estruturação e do funcionamento das redes neurais que, em tese, dão suporte físico à GU. Modelos *seletivos* de aquisição linguística, como o de Lightfoot (1989) e Anderson & Lightfoot (2000), baseiam-se em uma arquitetura cerebral em que os estímulos ambientais relevantes, ao serem analisados por uma rede de neurônios especialmente dedicada ao processamento da linguagem, selecionam e fixam determinadas conexões sinápticas (e as diretrizes linguísticas abstratas

intellectus ipse” (“exceto o próprio intelecto”) – transformaria este axioma empirista em um mote racionalista (cf. GUIMARÃES, 2017: Cap. 5). (Texto original: “All ideas come from sensation or reflection. Let us then suppose the mind to be, as we say, white paper, void of all characters, without any ideas: – How comes it to be furnished? Whence comes it by that vast store which the busy and boundless fancy of man has painted on it with an almost endless variety? Whence has it all the materials of reason and knowledge? To this I answer, in one word, from *experience*. In that all our knowledge is founded; and from that it ultimately derives itself.”).

que delas emergem). Assim, adquirir uma língua envolve amplificar e fixar certas sequências de sinapses e eliminar outras – todas elas já pré-estabelecidas pela estrutura da GU (o que corresponde a uma espécie de seleção neuronal darwiniana, na medida em que conexões sinápticas recorrentes e persistentes reduzem a superabundância conectiva das células do cérebro em desenvolvimento [DEAKON, 2000]). Este modelo seletivo de funcionamento biológico pode ser observado em outros dispositivos fisiológicos, como é o caso do processamento das informações luminosas realizado pelo córtex visual – em que os estímulos visuais advindos da experiência afetam diferencialmente a atividade de neurônios especializados na interpretação do sinal luminoso, favorecendo certas conexões sinápticas em detrimento de outras, que permanecem inativas ou acabam se degenerando –, e como também é o caso da amplificação imunológica que determinadas linhagens de anticorpos (proteínas) sofrem ao serem selecionadas de modo específico por antígenos: ao invés de sintetizar “do zero” proteínas capazes de desativar determinadas substâncias antigênicas, o sistema imunológico amplifica a ação de anticorpos já pré-existent no organismo, diminuindo o tempo de resposta do sistema de defesa (a esse respeito, cf. o texto *The generative grammar of the immune system* (1985), de Niels K. Jerne, vencedor do Prêmio Nobel de Medicina e Fisiologia de 1984. Nele, o autor explica como o modelo gramatical seletivo proposto por Chomsky – segundo o qual os estímulos linguísticos ambientais apenas selecionam/especificam princípios gramaticais inatos – foi um dos *insights* para a descoberta de um dos mecanismos de amplificação do sistema imunológico: “*Parece, à primeira vista, que a resposta do sistema imunológico a uma sentença [de aminoácidos] apresentada por uma molécula proteica invasora consiste simplesmente em selecionar, entre seu enorme repertório de anticorpos pré-formados, uma imagem espelhada adequada de parte dessa sentença antigênica*” [JERNE, 1985: 447]⁵³. Assim, a resposta eficiente do sistema imune a um novo antígeno invasor deriva do emparelhamento parcial de anticorpos já construídos, que possuem trechos “espelhados” capazes de se conectar a determinadas partes do antígeno, que é inativado sem que seja necessária a síntese de um anticorpo novo).

Por outro lado, modelos *instrucionais* de aquisição linguística se fundamentam na existência de uma base neuronal polivalente que é capaz de aprender diretamente através dos estímulos sensoriais, sem a necessidade de uma matriz cognitiva pré-configurada: “[...] De

⁵³ “It would seem, at first sight, that the immune response to a sentence presented by an invading protein molecule is merely to select, from amongst its enormous preformed antibody repertoire, a suitable mirror image of part of this antigenic sentence.”

acordo com uma teoria instrucional, um sinal externo transfere sua característica ao sistema que o recebe, instruindo o que é essencialmente um sistema nervoso plástico e modificável.” (LIGHTFOOT, 1989: 321)⁵⁴.

Em suma, em um modelo seletivo de aquisição, restrições e princípios estruturantes geneticamente determinados filtram e organizam os dados linguísticos captados pelos órgãos dos sentidos, gerando uma gramática interna a partir de uma base neuronal pré-programada para tal fim; em um modelo instrucional, estímulos linguísticos “ensinam” redes de neurônios pluripotentes a construir estruturas gramaticais unicamente a partir das informações contidas nos dados. Em termos lógico-matemáticos, quanto mais qualificado, especializado e seletivo é o sistema cognitivo de processamento linguístico, menos qualificados podem ser os dados advindos da experiência concreta do indivíduo em sua comunidade de fala, na medida em que o *input* funciona “apenas” como estabilizador de valores pré-configurados (cf. 11a); em sistemas instrucionais, ao contrário, quanto menor é o grau de especialização do aparelho de aquisição, mais qualificados devem ser os dados linguísticos extraídos do ambiente, a fim de dar conta de toda a complexidade alcançada pela gramática final do falante (cf. 11b).

$$(11) \quad \begin{array}{l} \text{a. } \uparrow|S_0| + \downarrow|D_a| \rightarrow |S_e| \\ \text{b. } \downarrow|S_0| + \uparrow|D_a| \rightarrow |S_e| \end{array}$$

Em (11), os módulos $|S_0|$ e $|D_a|$ indicam, respectivamente, a qualidade do sistema cognitivo inicial e a qualidade do conjunto de dados (entendendo *qualidade de X* como o grau de sobreposição (ou de congruência) entre *X* e o sistema final $|S_e|$). Note que estamos assumindo neste ponto uma visão monista entre cérebro e mente, em que S_0 corresponde tanto ao estágio cognitivo inicial da faculdade da linguagem (GU) quanto à sua base neurofisiológica. Esta unificação não é isenta de desafios, mas se estamos interessados em discutir as bases genéticas da faculdade da linguagem, precisamos ir em direção às propriedades materiais que dão suporte a essa faculdade – afinal, genes neuromodeladores, antes de regularem indiretamente propriedades cognitivas como a linguagem, primariamente

⁵⁴ “Under an instructive theory, an outside signal imparts its character to the system that receives it, instructing what is essentially a plastic and modifiable nervous system.”

codificam proteínas e são responsáveis pela construção da substância física de nossos cérebros⁵⁵.

Sistemas neurofisiológicos e cognitivos altamente organizados como os de (11a) circunscrevem módulos especializados em determinadas tarefas biológicas de processamento e representam verdadeiros órgãos mentais; sistemas não especializados como os de (11b) se baseiam em um cérebro maleável que é moldado pela experiência. Apesar da inegável plasticidade do cérebro humano – que é capaz de se reconfigurar e de recuperar parcialmente funções cognitivas afetadas por situações extremas de desestabilização, como acidentes vasculares cerebrais e neoplasias –, o processamento linguístico e o desenvolvimento de uma gramática interna são tarefas tão singulares que parece inviável biologicamente que tais funções especializadas não sejam executadas por um módulo – quer ele esteja instalado em um núcleo cerebral circunscrito, quer ele se distribua por uma rede de núcleos processadores interligados – especificamente voltado à linguagem:

[...] um órgão computacional com uma estrutura especialmente adaptada às demandas desse domínio específico. Quer esse órgão resida em uma parte altamente localizada do cérebro, quer ele surja de interconexões linguísticas específicas de diversos módulos cerebrais de processamento de dados, isso é irrelevante quanto a ele constituir um órgão distinto ou não. Alguns órgãos estão localizados em regiões específicas (por exemplo, os rins), enquanto outros se ramificam por vários lugares (por exemplo, o sistema circulatório). A característica essencial de um órgão é que ele tem uma função distinta da função de outros órgãos e uma estrutura

⁵⁵ O problema da relação entre corpo/cérebro (C) e mente (M) se organiza em torno de duas frentes filosófico-científicas: o naturalismo metafísico, em que os fenômenos abstratos de M estão amalgamados aos processos biofísicos que os geram em C, e o dualismo ontológico, em que C e M constituem entidades disjuntivas mutuamente irreduzíveis. “Subscrever ao naturalismo metafísico significa subscrever ao monismo fisicalista, que contrasta com o dualismo ontológico entre corpo e mente defendido por Descartes. Subscrever ao dualismo ontológico é supor que as entidades mentais não são entidades físicas, porque as primeiras não são redutíveis às últimas. Subscrever ao monismo fisicalista é assumir que todos os processos químicos, biológicos, psicológicos, linguísticos ou culturais são processos físicos que obedecem às leis fundamentais da física. Para um fisicalista, o problema ontológico da relação entre corpo e mente é descobrir como identificar a última com o primeiro por meio de uma redução ontológica.” (JACOB, 2010: 219). (Texto original: “To subscribe to metaphysical naturalism is to subscribe to physicalist monism, which stands in contrast to the ontological dualism between body and mind advocated by Descartes. To subscribe to ontological dualism is to suppose that mental entities are not physical entities, because the former are not reducible to the latter. To subscribe to physicalist monism is to assume that all chemical, biological, psychological, linguistic, or cultural processes are physical processes obeying the fundamental laws of physics. For a physicalist, the ontological problem of the relation between body and mind is to find out how to identify the latter with the former via an ontological reduction”).

adequada a essa função – uma estrutura que lhe permite realizar seu trabalho. Embora Chomsky (1975) estivesse lidando especificamente com a linguagem, ele claramente entendeu que sua concepção computacional de aprendizagem implicava que outras formas de aprendizagem também deveriam ser mediadas por órgãos de aprendizagem adaptados ao processamento de problemas específicos. Do ponto de vista computacional, a noção de um processo de aprendizagem de caráter geral (por exemplo, aprendizagem associativa), não faz mais sentido do que a noção de um órgão sensorial de função geral – uma protuberância no meio da testa cuja função é detectar e sentir coisas. Esse órgão, obviamente, não existe, porque captar informações de diferentes tipos de estímulos – luminosos, sonoros, químicos, mecânicos e assim por diante – requer órgãos com estruturas moldadas pelas propriedades específicas dos estímulos que processam. A estrutura de um olho – incluindo o circuito neuronal da retina e os circuitos além dela – reflete em detalhes apurados as leis da ótica e está adaptada às exigências de extração de informações sobre o mundo a partir do processamento da luz refletida. O mesmo vale para o ouvido, onde as demandas de extração de informação dos sons ditam as muitas propriedades distintivas dos órgãos auditivos. Vemos com os olhos e ouvimos com os ouvidos – em vez de sentir através de um órgão sensorial de caráter geral – porque a percepção requer órgãos sensoriais com estruturas especificamente adaptadas a cada modalidade. Chomsky percebeu que o mesmo deve ser verdade para a aprendizagem. Aprender coisas diferentes sobre o mundo a partir de diferentes tipos de experiência requer computações adaptadas tanto ao que deve ser aprendido quanto ao tipo de experiência a partir da qual esse aprendizado é realizado. Portanto, deve haver órgãos de aprendizagem específicos para determinados tipos de tarefas, com estruturas adaptadas tanto ao que eles devem extrair da experiência quanto ao tipo de experiência da qual eles devem extrair as informações. Por razões computacionais, pode-se esperar que os órgãos de aprendizagem difiram entre as diferentes espécies de animais, assim como ocorre com os órgãos sensoriais. As víboras detectam a radiação infravermelha, enquanto nós não, porque elas têm um órgão sensorial que nos falta. Nós aprendemos línguas e os chimpanzés não porque temos um órgão especialmente moldado para essa aprendizagem, algo que eles não têm (GALLISTEL, 2010: 194-195)⁵⁶.

⁵⁶ “[...] a computational organ with a structure tailored to the demands of this particular domain. Whether this organ resides in a highly localized part of the brain or arises from a language-specific interconnection of diverse data-processing modules in the brain is irrelevant to whether it constitutes a distinct organ or not. Some organs are localized (for example, the kidney) while others ramify everywhere (for example, the circulatory system). The essential feature of an organ is that it has a function distinct from the function of other organs and a structure suited to that function, a structure that makes it possible for it to do its job. Although Chomsky (1975) had language foremost in mind, he clearly understood that his computational conception of learning implied that other forms of learning must likewise be mediated by problem-specific learning organs. From a computational point of view, the notion of a general purpose learning process (for example, associative learning), makes no more sense than the notion of a general purpose sensing organ – a bump in the middle of the forehead whose function is to sense things. There is no such bump, because picking up information from different kinds of

A hipótese de construção de uma gramática interna a partir de um sistema não pré-configurado ($\downarrow|S_0|$), portanto, precisa lidar com o desafio de explicar como mecanismos cognitivos gerais (de memorização, de analogia, de indução e de repetição/imitação, por exemplo) são capazes de processar adequadamente estruturas de natureza ontológica única no universo, extraindo delas padrões e internalizando em *todas* as crianças de uma mesma comunidade de fala, ao longo do *mesmo* período crítico de aquisição, o *mesmo* conjunto essencial de regras e de restrições a partir de dados linguísticos compatíveis com inúmeras outras possibilidades. Na próxima seção, discutiremos os limites de um desses mecanismos – o indutivo – quando confrontado com determinados fenômenos de aquisição linguística.

3.2. ÓRGÃO DA LINGUAGEM VS. MECANISMOS GERAIS DA INTELIGÊNCIA

Considere o seguinte conjunto de sentenças do inglês⁵⁷:

- (12) a. Robots that can run hunt humans.
 b. Robots that run can hunt humans.
 c. Can robots that [_]₁ run [_]₂ hunt humans?

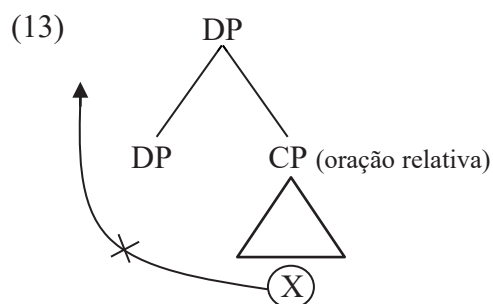
Em (12c), o fronteamento do auxiliar (ou modal) *can* dá origem a uma interrogativa polar que solicita como resposta a confirmação (“sim”) ou a negação (“não”) do valor de

stimuli – light, sound, chemical, mechanical, and so on – requires organs with structures shaped by the specific properties of the stimuli they process. The structure of an eye – including the neural circuitry in the retina and beyond – reflects in exquisite detail the laws of optics and the exigencies of extracting information about the world from reflected light. The same is true for the ear, where the exigencies of extracting information from emitted sounds dictates the many distinctive features of auditory organs. We see with eyes and hear with ears – rather than sensing through a general purpose sense organ – because sensing requires organs with modality-specific structure. Chomsky realized that the same must be true for learning. Learning different things about the world from different kinds of experience requires computations tailored both to what is to be learned and to the kind of experience from which it is to be learned. Therefore, there must be task-specific learning organs, with structures tailored both to what they are supposed to extract from experience and to the kind of experience from which they are to extract it. For computational reasons, learning organs may be expected to differ between species of animals, just as do sensory organs. Pit vipers sense infrared radiation, whereas we do not, because they have a sensory organ, which we lack. We learn languages, whereas chimpanzees do not, because we have a language-learning organ, which they lack.”

⁵⁷ Os pontos cruciais desta análise seguem a linha argumentativa de Chomsky (2011).

verdade da proposição subjacente à pergunta. Mas qual seria esta proposição? Qualquer falante nativo do inglês *sabe* que a interrogativa (12c) só pode ser derivada da sentença (12b) e que o “desejo” de quem a constrói é saber se robôs que correm podem (ou não) caçar humanos e não se os robôs podem (ou não) correr – ou seja, todo falante do inglês sabe que o auxiliar *can* se relaciona com a posição de base [_]₂ e com o complexo verbal *hunt humans* e não com a posição de base [_]₁ e o verbo *run*. Ora, se o processamento da sentença (12c) fosse orientado por um mecanismo cognitivo geral de economia centrado na facilidade de computação, não seria esperado que um falante nativo, ao invés de relacionar o auxiliar frontado *can* ao verbo *hunt*, o associasse ao verbo *run*, dada a proximidade linear existente entre os dois? Por que, afinal, o sistema gramatical do falante adota como critério de processamento a *distância estrutural mínima* – que relaciona (12c) a (12b) – e não a *distância linear mínima* – que relaciona virtualmente (12c) a (12a)? Se a gramática interna do falante fosse orientada por princípios cognitivos gerais de economia procedural, a computação da distância linear em (12c) (bem mais simples) deveria se sobrepor ao princípio de dependência estrutural (bem mais complexo), o que não ocorre.

A posição adequada de interpretação do modal *can*, em (12c), também deriva das fortes restrições que pesam sobre a interpretação deste elemento na posição [_]₁, já que, neste caso, o modal teria que ter sido extraído da oração relativa de (12a) (*robots* [_{CP} *that can run*] *hunt humans*). No entanto, desde Ross (1967), sabemos que as orações relativas se configuram como ilhas estruturais das quais (canonicamente) nenhum elemento pode ser extraído através de regras transformacionais. Assim, a associação do modal *can* à posição [_]₁ representaria a violação de uma imposição restritiva (teoricamente) universal – a *restrição de ilha adjunto* – que proíbe, entre outras coisas, a extração de um constituinte ilhado em uma oração relativa (cf. 13).



(RAPOSO, 1992: 395)

Sem um sistema cognitivo inicial enviesado gramaticalmente ($\uparrow|S_0|$), como explicar fenômenos como a dependência de estrutura e as restrições de ilha? Como vimos em (12c), se as diretrizes de processamento linguístico derivassem de princípios cognitivos gerais de economia computacional – que buscam sempre as soluções mais simples de interpretação –, as línguas humanas estariam condicionadas à linearidade das construções (e uma sentença como *o padre entrou na sala de lingerie* teria apenas uma interpretação – aquela que associa *lingerie* à sala e não ao padre). Qualquer teoria linguística que se alicerce, em maior ou menor grau, na existência de uma inteligência geral capaz de executar diferentes tarefas – entre elas, adquirir uma língua – precisa demonstrar em quais outros domínios cognitivos podemos encontrar operações semelhantes àsquelas encontradas durante o processamento da linguagem. Que mecanismo geral da inteligência seria capaz, por exemplo, de delimitar com precisão as ilhas de (12), estipulando as posições das quais o modal *can* pode ou não ser extraído? Será que uma criança em fase de aquisição poderia chegar a esta regra proibitiva através de operações indutivas? Saltos indutivos permitem que uma propriedade *p* seja generalizada após a verificação de que *p* é válida em um número *n* de casos particulares. Assim, se durante toda a sua existência um observador tem a experiência de ver o Sol despontando no horizonte após um período de escuridão, é provável que ele execute um salto indutivo e que crie a convicção de que “o Sol nasce todos os dias de manhã”, o que inclui as auroras do passado, do presente e do futuro.

Perceba que este tipo de cálculo mental – que generaliza uma propriedade *p* (o nascimento do Sol) a partir de um número limitado de eventos experienciados – não deriva de um raciocínio lógico, uma vez que não é possível inferir logicamente o futuro do passado: milhares de manhãs com sol não são suficientes para garantir o sol da manhã seguinte. Talvez hoje, com os conhecimentos astronômicos de que dispomos, possamos dizer mais convictamente que o Sol nascerá amanhã de manhã: afinal, a não ser que um evento astronômico catastrófico – como um buraco-negro devorando o sistema solar de uma hora para outra – modifique os movimentos relativos da Terra e do Sol, o nosso planeta continuará girando em torno de seu próprio eixo e em torno de nossa estrela, garantindo que, após a noite de hoje, experienciemos o dia de amanhã. Durante milênios, no entanto, nossos antepassados confiaram única e exclusivamente no raciocínio indutivo (acrescido de uma ponta de esperança) para calcular o sol da manhã seguinte. Foi este tipo de operação mental que deu à nossa espécie a capacidade de prever eventos e de planejar ações futuras. A indução, entretanto, sempre traz consigo um componente de incerteza. Mesmo com milhares de auroras

acumuladas em suas experiências, que tipo de garantia estes primeiros homens tinham de que o Sol viria após a madrugada? Quem poderia garantir que um poder desconhecido (quem sabe um deus) não os manteria em noite eterna em algum momento? No tempo em que os homens olhavam para o céu, talvez cada alvorada fosse celebrada como uma dádiva. Quantas manhãs ensolaradas foram necessárias para que nós realizássemos o salto (espetacular, do ponto de vista evolutivo) que permitiu irmos do “sol das manhãs vividas” em direção ao “sol de todas as manhãs”? Quantas ocorrências de um mesmo evento, afinal, são necessárias para que o salto indutivo seja realizado com segurança?

Imagine que exista um silo com dois trilhões ($2 \cdot 10^{12}$) de sementes brancas e com apenas uma semente verde. Imagine também que exista alguém disposto a retirar todas as sementes desse silo, uma de cada vez. Começada a tarefa, é improvável que essa pessoa, mesmo que já tenha retirado milhões de sementes brancas do silo, realize um salto indutivo e chegue a uma conclusão do tipo “todas as sementes deste silo são brancas”. Isto se dá porque nosso conhecimento de mundo (que inclui informações sobre silos e sobre a diversidade das sementes) impede que esta generalização seja feita. No entanto, imagine agora que esse silo esteja instalado em uma comunidade que todos os dias pela manhã envia um de seus moradores para acionar um mecanismo no silo a fim de receber dele uma única semente, que será transformada em chá para ser consumida por toda a comunidade em um ritual que se repete há milhões de anos (lembre-se de que estamos no reino da imaginação. Os agrupamentos mais antigos de *homo sapiens* datam de aproximadamente 200.000 anos – cf. BROWN *et al.*, 2012). A probabilidade de ocorrer o evento “semente verde” em uma manhã é dada pelo quociente entre o número de casos favoráveis (n (semente verde) = 1) e o número de casos possíveis (total de sementes no silo = T), i.e., P (semente verde) = $1 / T$. Assim, a probabilidade de a primeira semente (a “semente primordial”) ser verde é igual a: P (semente verde) = $1 / 10^{12}$, um número incrivelmente pequeno.

À medida que o tempo passa e a quantidade de sementes do silo diminui, aumenta a probabilidade de que o evento “semente verde” ocorra ($\uparrow P$ (semente verde) = $1 / \downarrow T$). No entanto, a quantidade de sementes no silo é tão grande que, mesmo que tenham se passado dois bilhões de dias desde o evento primordial (um intervalo de tempo que corresponde à colheita de dois bilhões de grãos), a chance de o evento “semente verde” ocorrer no ritual do dia seguinte continua extremamente pequena (P (semente verde) = $1 / (2 \cdot 10^{12} - 2 \cdot 10^9) = 1 / 1,998 \cdot 10^{12}$). Probabilisticamente, são as sementes brancas que constroem a realidade. Se esta comunidade é orientada por um raciocínio indutivo semelhante ao nosso e se o “presente do

silo” é considerado um fenômeno da natureza por esse grupo, é provável que, agora, a proposição “todas as sementes do silo são brancas” seja assumida como verdadeira pelos membros desse grupo (com exceção dos mais céticos, que provavelmente organizarão uma expedição para explorar o silo), uma vez que milhões de fenômenos concretos particulares fornecem um substrato de experiência que permite a construção de saltos em direção ao geral e ao abstrato. No entanto, nós sabemos que há uma semente verde no silo – e que um dia esta singularidade virá à tona, provavelmente derramando terror no coração da comunidade. Este exemplo – em que temos consciência total de todos os eventos possíveis – deixa claro, mais uma vez, que a indução não deriva de um raciocínio lógico: não é possível inferir logicamente uma propriedade da totalidade a partir de propriedades identificadas em instâncias particulares: 1 bilhão de sementes brancas recolhidas de um silo não garantem por si só que a próxima semente será da mesma cor. Hoje, sabemos que, pela idade da Terra, o Sol já nasceu aproximadamente 1,7 trilhão de vezes em nosso planeta. Uma quantidade, sem dúvida, astronômica de eventos. Mas que garantia temos de que o Sol virá depois desta noite? Não será amanhã o dia em que o silo cósmico nos dará a semente fatal?

Uma resposta adequada às perguntas que fecham o parágrafo anterior deriva do modo como, hoje, processamos as auroras. Como já dissemos anteriormente, com o conhecimento de que dispomos atualmente, podemos assegurar que o Sol despontará no horizonte amanhã cedo não apenas porque nosso instinto de indução nos leva a esta forte convicção, mas principalmente porque as Leis da Física assim o garantem. De fato, hoje, podemos calcular com precisão as posições relativas que o Sol e a Terra estabelecem entre si em função do tempo – o que significa não apenas calcular a alvorada de amanhã, mas também as do passado profundo e as do futuro distante. Modelos matemáticos podem estabelecer com exatidão em que ponto do sistema solar o nosso planeta estará daqui a cem mil anos ou em qual ponto do horizonte da Filadélfia o Sol nasceu em 07 de dezembro de 1928. Mesmo que um evento cósmico desconhecido (como um meteoro pulverizando a Terra) possa acabar com qualquer predição matemática sobre as auroras em função do tempo, o ponto que queremos destacar aqui é que não é mais a indução pura e simples que nos permite dizer que o dia de amanhã existirá – mas sim o nosso conhecimento sobre as leis naturais que regem a harmonia do Universo.

Voltemos à nossa pergunta: pode uma criança adquirir regras gramaticais proibitivas através de operações indutivas? Um fato interessante (e um tanto óbvio) sobre as línguas é que as restrições gramaticais não são explícitas – uma sentença como *robots that can run hunt*

humans não vem acompanhada da informação de que o modal *can*, nesse contexto, não pode ser deslocado para o início da construção a fim de formar uma interrogativa. Ora, sem esta proibição explícita etiquetada às sentenças e considerando que o *infans* lida o tempo todo com interrogativas fronteadas por *can*, não seria esperado que uma criança adquirindo a língua inglesa ao menos tentasse indutivamente arrancar o modal de dentro da oração relativa, colocando-o à frente da sentença para formar uma pergunta? O fato é que ela não o faz. E não o faz sistematicamente, indo muitas vezes na contramão do que seria esperado diante das evidências positivas do seu *input*. Consideremos as seguintes trincas de construções do PB:

14. a) A Maria viu o gato comer um torresmo.
 b) A Maria viu o gato comer o quê?
 c) O que_i a Maria viu o gato comer _____i?

15. a) A Maria disse que o gato comeu um torresmo.
 b) A Maria disse que o gato comeu o quê?
 c) O que_i a Maria disse que o gato comeu _____i?

16. a) A Maria viu o gato que comeu um torresmo.
 b) A Maria viu o gato que comeu o quê?
 c) *O que_i a Maria viu o gato que comeu _____i?

As *interrogativas QU* (também chamadas de *interrogativas de conteúdo*) são aquelas em que um pronome QU (*quem, o quê, como, quando, onde* etc.) dispara um processo de varredura informacional que procura estabelecer o valor adequado de uma variável *x*. Em uma sentença como (i) *Quem mexeu no meu queijo?*, por exemplo, o pronome *quem* aciona um processo de mapeamento do mundo que busca determinar com precisão o valor referencial da variável *x*, de modo que a proposição *x mexeu no meu queijo* seja verdadeira. Os pronomes interrogativos QU podem aparecer em sua posição de base (delineando aquilo que chamamos de interrogativas *in situ* (cf. 14b; 15b; 16b)) ou podem ser deslocados para o início da sentença (cf. 14c; 15c).

Vamos assumir que o *input* de qualquer criança adquirindo o PB possui sentenças estruturalmente semelhantes às dos exemplos 14-16, excetuando-se a construção (16c), que constitui a nossa evidência negativa. Dadas as derivações (14b → 14c) e (15b → 15c), seria esperado que uma criança guiada pelo espírito indutivo generalizasse a regra de fronteamto de sintagmas interrogativos e transformasse (16b) em (16c). Mas, como já dissemos, ela não o

faz. Ora, se estamos assumindo que esta criança não nasce pré-programada com regras restritivas (ou de qualquer espécie), então precisamos explicar como ela descobriu e internalizou, através da análise indutiva dos dados, a regra que proíbe a extração de pronomes QU do domínio de orações relativas. O problema é que os dados a que a criança tem acesso são compatíveis com inúmeras gramáticas e, no limite, com infinitas regras de restrição. Sem um módulo especificamente linguístico que guie a criança durante o processamento do *input*, ela pode construir um número potencialmente alto (e virtualmente incomputável) de hipóteses que são compatíveis com sua experiência linguística concreta e que podem, em tese, explicar por que sentenças como (16b) não podem ser transformadas em construções como (16c). Apenas a título de ilustração, vamos imaginar algumas delas: (i) “O fronteamento de um sintagma interrogativo não pode ocorrer quando o sujeito da oração encaixada está separado do pronome QU por x vocábulos”; (ii) “O fronteamento de um sintagma interrogativo não pode ocorrer quando a sentença possui y sílabas tônicas”; (iii) “O fronteamento de um sintagma interrogativo não pode ocorrer quando o falante A se dirige a um interlocutor B no ambiente X”; etc.

Talvez as hipóteses de restrição acima pareçam esdrúxulas para nós que já conhecemos a natureza gramatical da restrição de ilha; para uma criança dotada de mecanismos indutivos de processamento linguístico e não enviesada gramaticalmente, entretanto, elas representam hipóteses plausíveis passíveis de serem testadas. O fato é que o número de hipóteses restritivas que podem emergir da indução é tão grande que não haveria tempo hábil para que uma criança testasse cada uma delas até descobrir a restrição correta dentro do período crítico de aquisição da língua. Ao final, teríamos crianças com diferentes restrições – i.e., com diferentes gramáticas – internalizadas em S_e . Pode-se eventualmente contra-argumentar que uma criança pode chegar, sim, à restrição de ilha através de operações indutivas. Talvez. Mas como explicar o fato de que *todas* as crianças internalizam a *mesma* restrição aproximadamente na *mesma* fase de maturação do cérebro? Além disso, o fato de (16c) não aparecer no *input* da criança não é suficiente para que ela conclua, via indução, que (16c) é proibido pelo sistema. Afinal, a experiência da criança com o mundo é limitada e pode ser que a ausência de (16c) em seu *input* seja um mero fruto do acaso: assim como mil

sementes brancas não dizem nada a respeito da existência ou da inexistência de uma semente verde no silo, mil ocorrências do tipo (16b), por si só, não garantem que (16c) não exista⁵⁸.

Mas e se ilhas adjunto (assim como as restrições que pesam sobre elas) forem condicionadas por leis naturais – talvez não tão estáveis quanto as leis que regem os movimentos dos astros, mas estáveis o suficiente para transmitirem certas diretrizes linguísticas de uma geração a outra via mecanismos hereditários? Se essa hipótese estiver correta, a restrição de ilha nada mais é do que uma restrição cognitiva (derivada de princípios abstratos inatos) que modula (ou limita) o poder de regras transformacionais (como a de frenteamento de sintagmas QU). Desta forma, a impossibilidade de se formarem construções interrogativas do tipo (16c) deriva de um limite gramatical fixado pelo bioprograma responsável pela estrutura *default* do módulo linguístico de nossa mente/cérebro e não de um aprendizado via indução. Perceba que o processamento adequado de estruturas linguísticas insulares pressupõe que a criança conheça e domine um certo conjunto de noções e de operações gramaticais fundamentais. A ilha de (16), por exemplo, faz parte de um NP complexo formado por adjunção (cf. *A Maria viu* [DP *o* [NP [NP *gato*] [CP *que comeu um torresmo*]]]), ao passo que as extrações de (14) e (15) ocorrem em orações encaixadas completivas, que não representam barreiras ao frenteamento de pronomes QU (cf. *O que_i a Maria viu* [_{InfP} *o gato comer t_i?*] e *O que_i a Maria disse* [CP *que o gato comeu t_i?*])). Assim, para que a criança processe corretamente as construções de 14-16, é indispensável que seu sistema computacional não apenas disponha de uma matriz de mapeamento de ilhas gramaticais, mas que seja capaz de, entre outras coisas, (i) delimitar constituintes complexos (sintagmas), (ii) diferenciar complementos de adjuntos e (iii) manipular as categorias abstratas (lexicais e funcionais) que subjazem a cada um dos itens que, combinados, formam as expressões linguísticas.

Não há palavra em qualquer língua humana conhecida que não pertença a uma categoria. O fato de que cada palavra pertence a uma categoria é a base do funcionamento da sintaxe. Embora ninguém saiba realmente quantas palavras os seres humanos sabem exatamente (as estimativas vão de 50.000 a 90.000), o número de categorias é muito menor. Comparado com o

⁵⁸ Para um estudo de caso envolvendo a aquisição de interrogativas QU no PB, cf. Sikansi, 1999. Além de discutir o contraste entre interrogativas *in situ* e frenteadas, a autora analisa o fenômeno da inversão verbo-sujeito (cf. *Onde o João foi?* vs. *Onde foi o João?*) e a presença/ausência de elementos expletivos (cf. *Onde (que) o João foi?* e *Onde (é que) o João foi?*) nas interrogativas que formam o *input* e o *output* de uma criança em fase de aquisição.

número de palavras existentes em uma língua, é realmente apenas um punhado. Agora, suponha que as regras sintáticas fossem específicas para cada palavra em inglês [ou em qualquer outra língua]. Então você precisaria de milhares e milhares de regras sintáticas, presumivelmente várias regras para cada palavra. Mas se as regras sintáticas se aplicarem apenas a categorias, o número de regras pode ser significativamente reduzido. [...] no final, não são as palavras que são os blocos de construção da sintaxe, mas sim as categorias (KOENEMAN & ZEIJLSTRA, 2017: 36-37)⁵⁹.

Desta forma, o mapeamento de ilhas adjunto por uma criança em fase de aquisição implica o agenciamento de uma série de noções e de procedimentos que, ao que tudo indica, também se encontram geneticamente determinados – sejam os de natureza exclusivamente gramatical (como as noções de *sintagma*, *complemento*, *adjunto*, *transformação* etc.), sejam os de natureza cognitiva mais geral (como o princípio de categorização, que, ao ser aplicado à linguagem, permite que itens lexicais ou funcionais com traços semelhantes possam ser agrupados em classes específicas, tais como *nomes*, *verbos*, *determinantes* etc.).

3.3. HIPÓTESE INATISTA FORTE VS. HIPÓTESE INATISTA FRACA

Assumir que noções gramaticais tão específicas como Restrições de Ilha e Princípios de Ligação possam fazer parte da estrutura cognitiva inata responsável pela aquisição de uma língua significa assumir que estes princípios se encontram inscritos no código genético de nossa espécie, o que, sem dúvida, é uma hipótese ousada (*hipótese inatista forte*). Afinal, que pressões evolutivas teriam selecionado princípios cognitivos que, aparentemente, não nos oferecem nenhum tipo de vantagem adaptativa? Organizar a realidade em categorias mentais (classificando os animais em “perigosos” vs. “não perigosos”, por exemplo), realizar inferências a partir de pistas (ligando as pegadas que foram vistas nas proximidades da comunidade à presença de um animal grande e potencialmente perigoso) e veicular ideias através de cadeias sonoras ou visuais (com as quais o grupo pode organizar um plano de

⁵⁹ “There is no word in any known human language that does not belong to a category. The fact that every word belongs to a category is a cornerstone of the way syntax works. Although nobody really knows how many words human beings know exactly (the estimates go from 50,000 to 90,000), the number of categories is much smaller. Compared to the number of words in a language, it’s really just a handful. Now, suppose that syntactic rules were specific for every word in English. Then you would need thousands and thousands of syntactic rules, presumably several rules for every individual word. But if syntactic rules only apply to categories, the number of rules can be significantly reduced. [...] What we see is that words in the end are not the building blocks of syntax, but categories are.”

defesa detalhado ou simplesmente conversar no fim do dia) são competências cognitivas que representam vantagens biológicas evidentes frente aos fatores de resistência do meio (como o predatismo) e, por isso, tendem a ser transmitidas de geração a geração, preservando os genes que as condicionam e aumentando sua frequência na população ao longo do tempo; por sua vez, é difícil imaginar quais foram as vantagens evolutivas que regras gramaticais com alto grau de especialização – como a que impede a extração de um constituinte de uma oração relativa – puderam oferecer à sobrevivência de nossa espécie. Se a faculdade da linguagem possui uma contraparte material biomolecular localizada no DNA humano, os genes que a condicionam estão submetidos aos mesmos princípios evolutivos que pesam sobre o restante do mundo vivo, o que inclui os mecanismos de seleção natural, que tendem a eliminar os alelos – e as características que deles derivam – que não favorecem a adaptação do indivíduo, conservando apenas aquelas sequências nucleotídicas que de alguma forma beneficiam o organismo. Assim, fenômenos cognitivos como as restrições de ilha e os princípios que regem a ligação de anáforas, pronomes e expressões-R, caso possuam realmente uma base genética, representam um mistério evolutivo, dados o seu caráter acessório e sua aparente “não funcionalidade” enquanto dispositivo de adaptação ao meio.

Entretanto, uma possível solução teórica para essa aparente contradição evolutiva – que conserva princípios cognitivos que não oferecem vantagens adaptativas imediatas ao organismo – pode ser encontrada no modo como o mecanismo de expressão gênica intracelular se relaciona com a gama de traços fenotípicos que emergem do sistema metabólico de processamento proteico. Um gene, ao ser transcrito e traduzido, gera como produto uma cadeia de aminoácidos (uma proteína ou parte dela) que constitui o seu *efeito primário* (ou *fenótipo molecular*). Este produto proteico, por sua vez, ao participar de uma série de vias bioquímicas metabólicas, acaba dando origem a uma ou mais características fenotípicas identificáveis que formam o *efeito secundário* do gene. Assim, é preciso diferenciar a ação que o gene exerce no nível celular do(s) efeito(s) que ele provoca no nível fenotípico secundário. Normalmente, associa-se a um gene uma única característica fenotípica, numa relação direta de 1 para 1. No entanto, não apenas uma característica pode ser determinada por vários genes – vide os casos de *interação gênica* e de *herança quantitativa*⁶⁰ –, como um único gene pode determinar um conjunto de características

⁶⁰ Existem dois tipos básicos de *interação gênica*: “a primeira categoria consiste em interações entre alelos de um mesmo *locus* – em termos mais amplos, variações de dominância [o *locus* é o “endereço cromossômico” de um gene e normalmente é utilizado como uma forma “neutra” de referi-lo, sem que seja necessário especificar

fenotípicas distintas. Na *pleiotropia*, o efeito primário de um gene – seu produto proteico –, ao participar de inúmeras vias metabólicas, determina uma gama de efeitos fenotípicos distintos. O mesmo gene que determina o tamanho das asas da mosca *Drosophila melanogaster*, por exemplo, determina o tamanho de suas pernas (LOPES & ROSSO, 2010:

sua forma alélica. A *dominância* se refere às relações que os alelos que ocupam o mesmo *locus* estabelecem entre si. Em organismos diploides – i.e., em organismos que possuem dois lotes de cromossomos ($2n$) – um alelo dominante (A) condiciona um determinado fenótipo mesmo quando se encontra em dose simples ($A_$). Já um alelo recessivo (a) só condiciona uma característica fenotípica quando se encontra em dose dupla, em indivíduos homozigóticos recessivos (aa)]. Embora essa informação não abranja a variedade de genes que afetam uma função, pode-se aprender muito sobre o papel de um gene ao considerar as interações alélicas. A segunda categoria consiste em interações entre dois ou mais *loci* [i.e., entre dois ou mais genes]. Tais interações revelam o número e os tipos de genes no programa global subjacente a uma função biológica particular.” (GRIFFITHS *et al.*, 2013: 182). Na *herança quantitativa* (ou *poligênica*), uma mesma característica fenotípica é determinada pelos efeitos cumulativos de dois ou mais genes, cada um deles colaborando com uma fração do fenótipo. “As características quantitativas [como a altura, o peso e a pressão arterial nos seres humanos] surgem a partir de dois fenômenos. Primeiro, muitas são poligênicas: são influenciadas por genes em vários *loci*. Se vários *loci* participarem, vários genótipos [i.e., várias configurações de alelos] são possíveis, cada um produzindo um fenótipo discretamente diferente. Segundo, as características quantitativas surgem quando os fatores ambientais afetam o fenótipo, [neste caso,] as diferenças ambientais fazem com que um único genótipo produza uma gama de fenótipos. A maioria das características que variam continuamente é poligênica e influenciada por fatores ambientais e estas características são chamadas de multifatoriais” (PIERCE, 2016: 996-997). A cor da pele dos seres humanos, por exemplo, é uma característica contínua (quantitativa) determinada pelo efeito cumulativo de um conjunto de genes associados à produção de melanina e também por fatores ambientais como a exposição maior ou menor aos raios solares. Funções cognitivas como a competência linguística humana emergem de circuitos neuronais determinados pela interação de inúmeros genes. “[...] nenhum gene funciona sozinho. Estruturas biológicas complexas – quer falemos do cérebro, do coração ou dos rins – são o produto das ações e das interações combinadas de muitos genes, não apenas de um. Uma razão pela qual não faz sentido correlacionar um único gene a um comportamento específico é que o circuito neuronal envolvido na produção de qualquer comportamento é muito mais complexo do aquele que poderia ser condicionado por qualquer gene. Não pode haver um único gene para a linguagem [...] assim como não pode haver um gene único para o ventrículo esquerdo do coração humano. Mesmo uma única célula cerebral – ou uma única célula cardíaca – é o produto de muitas proteínas e, portanto, de muitos genes.” (MARCUS, 2004: 80. No original: “[...] no gene works on its own. Complex biological structures – whether we speak of the brain or of hearts or kidneys – are the product of the concerted actions and interactions of many genes, not just one. One reason that it makes no sense to talk about a gene “for” a particular behavior is that the neural circuitry involved in producing any given behavior is far more complex than any one gene. There can no more be a single gene for language [...] than there can be for the left ventricle of a human heart. Even a single brain cell – or a single heart cell – is the product of many proteins and hence many genes.”).

345). Muitos dos *erros inatos do metabolismo* consistem em efeitos pleiotrópicos causados por mutações que afetam um único gene. Algumas destas síndromes genéticas pleiotrópicas são mostradas na tabela da Figura 3.3, que se refere à espécie humana.

Condição clínica	Herança autossômica (localização cromossômica)	Efeito primário	Efeitos secundários
Doença de Waardenburg, tipo I	Dominante (2q)	Alteração no controle da transcrição por mutação com perda de função do gene <i>PAX3</i>	Mecha branca no cabelo, surdez neurossensitiva congênita, nariz com ponte alta e larga, deslocamento lateral dos ângulos médios dos olhos, heterocromia e hipoplasia da íris
Fenilcetonúria	Recessiva (12q)	Deficiência da enzima fenilalanina-hidroxilase	Deficiência mental, excreção de fenilcetonas na urina, pigmentação clara e odor de mofo
Síndrome de Marfan	Dominante (15q)	Defeito nas fibras elásticas do tecido conectivo	Anomalias esqueléticas, com extremidades alongadas, deslocamento do cristalino e anomalias cardiovasculares
Porfíria variegada	Dominante (14q)	Defeito no metabolismo da porfírina, componente da hemoglobina	Acúmulo de porfírina na urina, dando-lhe tom vermelho-escuro, dor abdominal, fraqueza muscular, cefaleias, insônia, problemas visuais que podem levar à cegueira, delírio e convulsões

FIGURA 3.3. Alguns exemplos de pleiotropia (um gene, vários efeitos). A herança autossômica se refere aos cromossomos não sexuais (autossomos). Fonte da imagem: BORGES-OSÓRIO & ROBINSON, 2013.

Desta forma, é possível que, em um cenário hipotético, alguns dos princípios da faculdade da linguagem se configurem como “efeitos colaterais” (exaptações) de um *cluster* de genes pleiotrópicos responsáveis pela determinação de funções biologicamente relevantes – o que explicaria a “sobrevivência” de tais princípios ao longo de nossa história evolutiva⁶¹. Imaginar que exista um gene R especialmente dedicado à noção abstrata de restrição de ilha, por exemplo, parece implausível do ponto de vista biológico. Mas se esse mesmo gene R estiver ligado a uma função (cognitiva ou não, linguística ou não) que represente uma vantagem ao organismo que a possui, a tendência é que esta sequência de DNA – e todos os produtos fenotípicos que dela derivam, incluindo os gramaticais – seja

⁶¹ *Exaptações* são adaptações biológicas que não surgiram por pressões seletivas relacionadas à sua função atual, mas sim por pressões evolutivas que incidiram sobre outras funções geneticamente relacionadas. Este, aliás, é um dos problemas relacionados à evolução de determinados aspectos da faculdade da linguagem. “[...] a questão da ‘continuidade *versus* exaptação’ gira em torno do problema de saber se a linguagem humana evoluiu por extensão gradual de sistemas de comunicação preexistentes, ou se aspectos importantes da linguagem foram exaptados de sua função adaptativa anterior (por exemplo, raciocínio espacial ou numérico; inteligência social maquiavélica [habilidade cognitiva relacionada à capacidade de socialização estratégica]; *tool-making* [competência associada à fabricação de ferramentas])” (HAUSER *et al.*, 2002: 1570). (Texto original: “[...] the “continuity versus exaptation” issue revolves around the problem of whether human language evolved by gradual extension of preexisting communication systems, or whether important aspects of language have been exapted away from their previous adaptive function (e.g., spatial or numerical reasoning, Machiavellian social scheming, tool-making”).

transmitida às gerações posteriores e se incorpore ao patrimônio genético da espécie. Um raciocínio semelhante pode ser utilizado para explicar por que determinados transtornos linguísticos parecem ser indissociáveis de outros transtornos cognitivos (sendo, inclusive, herdados em conjunto). A coocorrência de déficits de natureza distinta pode derivar da *“existência de uma base genética parcialmente comum para dois ou mais tipos de capacidades cognitivas”* (BURRACO, 2012: 45). Assim, se essa base genética parcialmente compartilhada sofre algum dano, todos os domínios cognitivos que são de alguma forma condicionados por ela podem ser afetados.

Como podemos ver, a versão forte do inatismo linguístico precisa lidar com problemas extremamente complexos quando confrontada com as leis do mundo natural. A partir do momento em que os princípios fundamentais da faculdade da linguagem passam a ser considerados parte intrínseca da biologia humana, apresentando uma contraparte física materializada na substância eletrofisiológica do cérebro e nos genes que a sintetizam, a GU passa a obedecer às leis da matéria e da genética. Se princípios como o de Restrição de Ilha são condicionados geneticamente, mutações nos genes que os condicionam devem provocar a degeneração do princípio em análise, ocasionando o aparecimento de déficits linguísticos de natureza diversa. Estes fenótipos desviantes podem, então, ser utilizados para o mapeamento, a identificação e o sequenciamento do(s) gene(s) associados a uma determinada diretriz da competência linguística. Assim, não deveríamos encontrar indivíduos capazes de realizar espontânea e consistentemente movimentos de extração de constituintes de ilhas gramaticais, gerando construções interpretáveis? Se a Restrição de Ilha faz parte do nosso genoma, não deveríamos encontrar crianças portadoras de mutações que lhes permitissem gerar expressões linguísticas não moduladas por esta proibição? A universalidade irrestrita da Restrição de Ilha – que muitas vezes é escolhida como cavalo de batalha quando queremos apresentar a ideia de que existem princípios que fazem parte da faculdade linguística de todos os seres humanos –, paradoxalmente, representa um problema de adequação teórica se a concebemos como objeto biolinguístico. Se ela de alguma forma se encontra codificada em nosso DNA, deveríamos encontrar fenótipos mutantes. Deveríamos encontrar exceções.

Problemas dessa natureza abriram espaço para um movimento de recuo na teoria, fazendo com que surgissem propostas mais moderadas de conceptualização da faculdade da linguagem, concebendo-a como uma disposição neurofisiológica determinada geneticamente, mas sem se comprometer com a delimitação, a descrição (e mesmo com a existência) de princípios abstratos inatos.

Diante da implausibilidade biológica do inatismo forte, estudiosos gerativistas mais moderados têm defendido o que entendemos como uma *versão fraca* da hipótese inatista. Tal como argumentam os psicólogos cognitivos Steven Pinker⁶² (da Universidade de Harvard) e Gary Marcus⁶³ (da Universidade de Nova York) e o linguista Paul Smolensky (da Universidade John Hopkins), a versão fraca do inatismo propõe que a faculdade da linguagem seja uma espécie de disposição biológica, algo como um conjunto de instruções genéticas que criam neurônios especializados em retirar informações linguísticas do ambiente a fim de construir uma língua-L. Essas instruções seriam uma espécie de *Mecanismo de Aquisição da Linguagem* (Language Acquisition Device – LAD, na sigla em inglês). [...] Na hipótese fraca, descarta-se a ideia de que todas as informações linguísticas que tornam a aquisição da linguagem possível estejam pré-detalhadas no genoma humano. Em vez disso, a hipótese fraca assume que os princípios linguísticos na verdade emergem durante o processo de aquisição da língua do ambiente, à medida que os neurônios especializados em linguagem formam sinapses em função dos estímulos da língua-E. Nesse sentido, para que o conhecimento linguístico possa emergir nos seres humanos, o mínimo suficiente e necessário é que os neurônios da linguagem sejam capazes de identificar e interpretar todos os detalhes da informação linguística a que a criança está exposta. Para tanto, basta que tais neurônios tenham hipersensibilidade à informação linguística, sendo capazes, portanto, de aprender com o sinal linguístico disponível no ambiente. (KENEDY, 2013: 78).

Deste modo, segundo a versão fraca do inatismo linguístico, diretrizes gramaticais como o Princípio C – que, como vimos, impede que uma expressão-R possa ser c-comandada por um elemento correferencial (cf. *Ele_i disse que [o capitão do navio]*_{i/j} pescou um tubarão*) – decorrem do processamento linguístico realizado por neurônios especializados (construção *a posteriori* do princípio) e não de regras *aprioristicamente* pré-configuradas na mente/cérebro (e no genoma) dos seres humanos.

RESUMO DO CAPÍTULO E CONSIDERAÇÕES FINAIS

Bebês humanos nascem geneticamente programados para adquirir a língua do ambiente sociointeracional em que vivem, apresentando um dispositivo inato de captação e processamento de informações linguísticas ambientais que serão utilizadas como valores de formatação de uma gramática interna. Diretrizes translinguísticas – tais como os Princípios de Ligação e as Restrições de Ilha – reforçam a hipótese da existência de uma base cognitiva especializada (a faculdade da linguagem) que integra a arquitetura cerebral/mental de todos os

⁶² PINKER, 1994: Cap. 10.

⁶³ MARCUS, 2004: Cap.7.

seres humanos e que, conseqüentemente, subjaz estruturalmente a todas as línguas existentes ou já extintas. Neste sentido, uma língua (do ponto de vista do indivíduo) é um fenótipo que emerge de um sistema neurológico/cognitivo pré-configurado, modulado por fatores ambientais (o *input* linguístico). Propostas centradas na existência de sistemas de aquisição linguística não especializados ($\downarrow|S_0|$) e em mecanismos gerais da inteligência (como a indução) não explicam satisfatoriamente como as crianças chegam (uniformemente) à internalização de uma gramática particular. Modelos que se apoiam em sistemas modulares de aquisição, com um bioprograma dedicado exclusivamente à linguagem ($\uparrow|S_0|$), explicam elegantemente como seres humanos bebês mapeiam a realidade de modo orientado, absorvendo e transformando sons (ou sinais visuais) carregados com traços linguísticos em uma gramática interior.

O arranjo harmônico entre forma e função atravessa todo o mundo vivo. Em um organismo, cada componente apresenta uma estrutura adequada à atividade que executa. O aparelho bucal mastigador de lagartas permite que elas se alimentem de folhas até o momento da metamorfose, quando então se transforma em uma tromba espiralada com a qual a borboleta pode sugar o néctar escondido no fundo das flores. As raízes filamentosas das plantas possibilitam que elas absorvam eficientemente água e nutrientes do solo; a geometria laminar de suas folhas permite que elas captem a luz solar. As fibras musculares robustas arranjadas em forma de câmaras dão aos corações humanos a capacidade de bombear sangue de forma ritmada para o resto do corpo. Como todos os outros órgãos, espera-se que as estruturas neurológicas de uma determinada espécie também tenham suas formas associadas às funções que realizam e que tenham sido moldadas evolutivamente pelas propriedades específicas dos estímulos que processam – como o são os circuitos neuronais que interpretam as informações luminosas captadas pelos olhos, por exemplo. Seria surpreendente, portanto, se o cérebro humano não possuísse um módulo funcional especialmente dedicado ao processamento de informação linguística, dada a natureza única deste tipo de estímulo. Assim como só conseguimos traduzir em cores uma determinada faixa do espectro eletromagnético porque possuímos um dispositivo cognitivo dedicado a tal tarefa de computação óptica, só conseguimos computar linguisticamente ondas sonoras e estímulos visuais porque o tecido neurológico e o programa mental que os processam foram moldados genética e evolutivamente para tal fim.

Hipóteses inatistas fracas apostam na existência de um mecanismo de aquisição da linguagem composto por neurônios especializados em processar informação linguística,

retirando dela os padrões que serão utilizados para construir caminhos sinápticos correspondentes a regras gramaticais. Hipóteses inatistas fortes consideram que, além de um módulo cerebral/mental especialmente dedicado à linguagem, determinados princípios gramaticais do programa linguístico inato de bebês já vêm configurados desde o momento em que o cérebro é construído, ainda na vida intrauterina. Em ambos os modelos, a faculdade da linguagem emerge como um componente cognitivo cujas diretrizes se encontram inscritas no código genético da espécie humana.

CAPÍTULO 4 – (ALGUNS) FATORES GENÉTICOS ASSOCIADOS À LINGUAGEM

Como a atividade gênica se expressa diferencialmente em função da atividade metabólica? Quais são os níveis de controle genético que subjazem ao órgão da linguagem? De que forma é possível mapear os genes envolvidos na resolução de puzzles linguísticos? Qual é a natureza bioquímica dos erros inatos do metabolismo e como alguns deles afetam determinadas funções da cognição linguística? O que são transtornos específicos de linguagem e quais são os mecanismos genéticos que os condicionam?

4.1. MODELOS E MÉTODOS DE INVESTIGAÇÃO

Em algum momento de nossas vidas (que poderíamos chamar de “instante inicial”), cada um de nós foi uma única célula que carregava em seu núcleo um arranjo aleatório de cromossomos transferidos por nossos pais. Nesse sistema inicial de 46 fragmentos de DNA estavam alocadas as informações que orientariam a montagem e o funcionamento de nossos organismos – definindo, por exemplo, de que modo, em que local e em qual período do desenvolvimento embrionário o nosso futuro coração e o nosso futuro cérebro seriam construídos e instalados. Após sofrer bilhões de divisões mitóticas, correspondentes a bilhões de replicações do material genético⁶⁴, esse sistema inicial totipotente se transformaria em um

⁶⁴ A velocidade e a precisão do mecanismo de replicação do DNA (que, em média, insere menos de um erro a cada 10 bilhões de nucleotídeos pareados) são características notáveis da engenharia celular de conservação informacional. “O tempo necessário para que a *Escherichia coli* [um procarionte, i.e., um ser unicelular cujo material genético não se encontra delimitado por um envoltório nuclear] replique seu cromossomo [único e circular] pode ser tão curto quanto 40 min. Portanto, seu genoma de cerca de cinco milhões de pares de bases tem de ser copiado a uma velocidade de cerca de 2.000 nucleotídeos por segundo. [...] O que é incrível nesse processo todo de replicação do DNA é que ele não sacrifica a velocidade pela acurácia. Como [esse mecanismo] pode manter tanto a velocidade como a acurácia, considerando a complexidade das reações na forquilha de replicação [o local onde ocorre a deselcoidização da dupla-hélice do DNA e a sua duplicação a partir das cadeias polinucleotídicas expostas que servem como molde para a síntese de novos filamentos (replicação semiconservativa)]? A resposta é que a DNA polimerase [uma enzima] faz parte de um grande complexo

aglomerado imenso de células que, apesar de terem se diferenciado e formado tecidos distintos, continuariam (ou melhor, continuam) portando o mesmo conjunto primordial de genes em seus núcleos. Cada tipo histológico, no entanto, exerce uma atividade celular específica e, por isso, possui uma configuração específica de genes que estão sendo expressos ou não. Um neurônio e uma célula muscular cardíaca, por exemplo, pela natureza das atividades que executam e pela fisiologia de seus metabolismos, possuem demandas distintas de produtos gênicos (i.e., de proteínas) e, portanto, apresentam conjuntos diferentes de genes ativos (ou seja, de trechos de DNA que estão sendo transcritos e traduzidos) em seus núcleos. Taxas diferenciais de expressão gênica também são corresponsáveis pela neuroplasticidade, *“cujo objetivo é a consecução de um ajuste o mais exato possível dos circuitos neuronais às demandas de processamento de informação geradas pelo meio externo”* (BURRACO, 2012: 71). Segundo Hofmann (2003), esta capacidade de ajuste e de remodelação neuronal pode corresponder a uma neuroplasticidade a curto prazo – em que a execução de uma tarefa cognitiva trivial aciona mudanças na atividade neuronal recorrendo-se apenas ao estoque polipeptídico já sintetizado pelos neurônios (modulação pós-traducional), sem que seja necessária a síntese de novos produtos proteicos – ou a uma plasticidade neuronal a longo prazo, em que mudanças profundas no ambiente neuroquímico (muitas delas tributárias do próprio processo de desenvolvimento neurológico do indivíduo) disparam vias metabólicas de processamento de sinal dentro dos neurônios, gerando reações intracelulares que promovem a modulação transcricional da atividade gênica, i.e., que promovem a expressão de novos genes e/ou a alteração dos níveis de expressão de genes já ativos, resultando na síntese de novas proteínas e na modificação/adaptação funcional de determinado circuito neuronal.

A extensão da remodelação morfofuncional envolvida na plasticidade neuronal depende da natureza e da intensidade das modificações ambientais que elicitam os diferentes processos moleculares e fisiológicos responsáveis por esta remodelação. Existem, necessariamente, modificações fisiológicas, estruturais e cognitivas de longo prazo, decorrentes de uma alteração do padrão de expressão gênica cerebral, que geralmente são causadas pelas diretrizes contidas no próprio plano de desenvolvimento ontogenético do organismo, pela sua trajetória de vida

‘nucleoproteico’ que coordena as atividades na forquilha de replicação. Esse complexo, chamado de replissomo, é um exemplo de ‘maquinário molecular’.” (GRIFFITHS *et al.*, 2013: 232). A complexidade e o tamanho do genoma dos organismos eucarióticos, cujo DNA, além de estar associado a proteínas (histonas), encontra-se distribuído em vários cromossomos, fazem com que a replicação eucariótica seja ainda mais impressionante, uma vez que o replissomo precisa desmontar os complexos formados por DNA e proteínas e coordenar a duplicação de múltiplos cromossomos, emparelhando bilhões de bases em um espaço de tempo relativamente curto.

e/ou pelo aprendizado. Pelo contrário, estímulos ambientais frequentes e/ou que requerem uma resposta imediata, traduzem-se em alterações na modulação do comportamento ou da aprendizagem, que, pela modificação da atividade ou da excitabilidade neuronais, ou por respostas de natureza endócrina [hormonal], provocam modificações a curto prazo na atividade cognitiva. Por sua vez, mudanças de curto prazo podem acabar causando modificações na expressão gênica e, portanto, podem levar a mudanças mais duradouras [...] (BURRACO, 2012: 74-76).

Somente um cérebro dotado de plasticidade pode dar conta de processar demandas e dados informacionais fluidos constantemente renovados. Tarefas triviais de computação linguística – como as que estão envolvidas na leitura desse parágrafo ou como as que estão relacionadas com a construção de declarações utilizadas no dia a dia – e tarefas complexas de aquisição – como aquelas envolvidas no mapeamento seletivo dos traços sonoros distintivos que, ao serem processados pela competência linguística de uma criança, formam o inventário mental de fonemas da língua que lhe serve de *input* – são *puzzles* que, para serem resolvidos, exigem o acionamento de sequências de sinapses que são até certo ponto inéditas no cérebro do indivíduo em que elas se realizam, fazendo com que o seu sistema neurológico seja remodelado a fim de proceder à execução adequada dessas tarefas. Essa remodelação neurofisiológica/cognitiva corresponde a uma taxa diferencial de expressão gênica que lhe permite a reformatação. Assim, as mudanças que ocorrem no padrão de expressão gênica cerebral em resposta aos estímulos variáveis do ambiente ou em resposta aos estímulos internos ao indivíduo podem ser utilizadas como uma forma de rastreamento dos genes que de algum modo estão envolvidos no processamento e na realização de uma determinada tarefa cognitiva – como, por exemplo, a aquisição de uma língua natural.

Aplicando os princípios gerais propostos por Hofmann (2003) para a caracterização da plasticidade fenotípica por meio da metodologia da genômica funcional [que consiste na identificação e na caracterização do conjunto de proteínas e/ou de moléculas de RNAm presentes em um tecido responsável por uma determinada função, o que permite detectar os genes que estão sendo expressos localmente], seria plausível propor a análise das modificações que ocorrem no padrão de ativação gênica nos centros cerebrais associados à linguagem conforme o organismo responde às mudanças produzidas no meio ambiente ou conforme responde ao próprio processo ontogenético. Nessa análise genética do surgimento da linguagem, os genes que apresentam variação nos níveis de expressão ao longo da ontogenia ou em resposta a alterações ocorridas no ambiente linguístico seriam aqueles direta ou indiretamente envolvidos no desenvolvimento e funcionamento do “órgão da linguagem” (BURRACO, 2012: 77).

A princípio, portanto, programas de investigação genômica como o de Hofmann (2003), que se baseiam na aferição comparativa de taxas diferenciais de produtos gênicos sintetizados durante a realização de tarefas cognitivas, permitem identificar quais genes entram em ação no momento em que o cérebro precisa processar/gerar linguagem – especialmente durante o período crítico de aquisição, quando o “órgão da linguagem”, que nessa fase corresponde a um sistema biogramatical relativamente indiferenciado (i.e., não totalmente configurado), precisa modelar os circuitos sinápticos que formatarão a gramática interna da criança. Modelos genômicos teóricos como o de Steven Pinker (1994) também partem da ideia de que a atividade gênica diferencial está na base do processamento gramatical (ou, em outras palavras, está na base da resolução de demandas/problemas linguísticos).

Tente ficar comigo nessa jornada neuromitológica: estamos começando a nos aproximar dos “genes da gramática”. As moléculas que guiam, conectam e preservam os neurônios são proteínas. Cada proteína é especificada por um gene, e um gene é uma sequência de bases da cadeia de DNA encontrada em um cromossomo. Um gene é ativado por “fatores de transcrição” e outras moléculas regulatórias - dispositivos que se prendem a uma sequência de bases em algum lugar de uma molécula de DNA e descompactam um trecho vizinho, permitindo que o gene seja transcrito em RNA, que é então traduzido em proteína. Geralmente, esses fatores regulatórios são proteínas, de tal modo que o processo de construção de um organismo é uma cascata intrincada de DNA produzindo proteínas, algumas das quais interagem com outro DNA para fazer mais proteínas, e assim por diante. Pequenas diferenças no ritmo de produção ou na quantidade de alguma proteína podem ter grandes efeitos no organismo que está sendo construído. Assim, um único gene raramente especifica alguma parte identificável de um organismo. Em vez disso, ele condiciona a liberação de alguma proteína em momentos específicos do desenvolvimento, contribuindo com um ingrediente em uma receita insondavelmente complexa, geralmente tendo algum efeito na moldagem de um conjunto de peças que também são afetadas por muitos outros genes. A construção do cérebro, em particular, tem uma relação complexa com os genes que a determinam. Uma molécula de superfície talvez não seja usada em um único circuito, mas em muitos, cada um guiado por uma combinação específica. Por exemplo, se houver três proteínas, X, Y e Z, que podem se fixar a uma membrana, um axônio pode se conectar a uma superfície que tem X e Y e não Z, e outro pode se conectar a uma superfície que tem Y e Z, mas não X. Os neurocientistas estimam que cerca de trinta mil genes, a maioria do genoma humano, são usados para construir o cérebro e o sistema nervoso. E tudo começa com uma única célula, o óvulo fertilizado. Ele contém duas cópias de cada cromossomo, uma advinda da mãe, uma do pai. Cada lote cromossômico parental foi originalmente montado nas gônadas dos pais pelo entrelaçamento

aleatório dos cromossomos dos dois avós⁶⁵. Agora chegamos a um ponto em que podemos definir o que seriam os genes gramaticais. Os genes da gramática seriam trechos de DNA que codificam proteínas (ou fatores de transcrição de proteínas) em determinados momentos e locais do cérebro, orientando, atraindo ou conectando neurônios em redes que, em combinação com as sinapses que ocorrem durante a aprendizagem, são necessários para computar a solução de algum problema gramatical (como escolher um afixo ou uma palavra). (PINKER, 1994: 321-322)⁶⁶.

No modelo de Pinker, portanto, a atividade linguística possui como contraparte uma atividade gênica neuronal que se expressa diferencialmente durante a execução de determinadas tarefas gramaticais. No período crítico de aquisição, além de serem ativados

⁶⁵ O princípio da segregação independente (segunda lei de Mendel) estabelece que genes em cromossomos diferentes se segregam de forma independente na meiose, combinando-se de maneira randômica nos gametas. Assim, se um *locus* A está localizado em um par de cromossomos homólogos e um *locus* B está localizado em outro par, um indivíduo duplo-heterozigoto para esses *loci* apresentará a configuração AaBb e formará quatro tipos de gametas quanto aos possíveis arranjos alélicos: AB, Ab, aB, ab, todos na mesma proporção.

⁶⁶ “Try to stay with me in this neuro-mythological quest: we are beginning to approach the “grammar genes.” The molecules that guide, connect, and preserve neurons are proteins. A protein is specified by a gene, and a gene is a sequence of bases in the DNA string found in a chromosome. A gene is turned on by “transcription factors” and other regulatory molecules—gadgets that latch on to a sequence of bases somewhere on a DNA molecule and unzip a neighboring stretch, allowing that gene to be transcribed into RNA, which is then translated into protein. Generally these regulatory factors are themselves proteins, so the process of building an organism is an intricate cascade of DNA making proteins, some of which interact with other DNA to make more proteins, and so on. Small differences in the timing or amount of some protein can have large effects on the organism being built. Thus a single gene rarely specifies some identifiable part of an organism. Instead, it specifies the release of some protein at specific times in development, an ingredient of an unfathomably complex recipe, usually having some effect in molding a suite of parts that are also affected by many other genes. Brain wiring in particular has a complex relationship to the genes that lay it down. A surface molecule may not be used in a single circuit but in many circuits, each guided by a specific combination. For example, if there are three proteins, X, Y, and Z, that can sit on a membrane, one axon might glue itself to a surface that has X and Y and not Z, and another might glue itself to a surface that has Y and Z but not X. Neuroscientists estimate that about thirty thousand genes, the majority of the human genome, are used to build the brain and nervous system. And it all begins with a single cell, the fertilized egg. It contains two copies of each chromosome, one from the mother, one from the father. Each parental chromosome was originally assembled in the parents' gonads by randomly splicing together parts of the chromosomes of the two grandparents. We have arrived at a point at which we can define what grammar genes would be. The grammar genes would be stretches of DNA that code for proteins, or trigger the transcription of proteins, in certain times and places in the brain, that guide, attract, or glue neurons into networks that, in combination with the synaptic tuning that takes place during learning, are necessary to compute the solution to some grammatical problem (like choosing an affix or a word).”

durante a (ou pela) atividade linguística, os genes que subjazem a essas tarefas, em conjunto com outras classes de genes, também podem ser expressos diferencialmente em função do próprio processo de desenvolvimento e maturação do órgão neurocognitivo responsável pela computação gramatical – i.e., em função de fatores ontogenéticos que controlam a atividade gênica, definindo quando, onde e em que quantidade ela ocorrerá. Assim, o programa genético responsável pela construção, pelo funcionamento e pela homeostase do órgão da linguagem apresentaria dois níveis básicos de ação gênica. Um primeiro nível seria formado pelos genes responsáveis pelo desenvolvimento histológico e pela organização sináptica dos núcleos linguísticos, incluindo os genes que regulam a atividade dos neurotransmissores (substâncias responsáveis pela transmissão de impulsos informacionais entre os neurônios) e os que controlam as interações entre os neurônios e as células da glia (tipos celulares que, em linhas gerais, têm a função de apoiar, nutrir ou proteger os neurônios⁶⁷). Nesse nível, portanto, estariam localizados os genes que determinam a estrutura geral do órgão da linguagem, condicionando seu desenvolvimento fisiológico “natural”. Na síndrome de Phelan-McDermid (ou síndrome de deleção do fragmento 22q13.3⁶⁸), por exemplo, uma mutação do gene

⁶⁷ “A glia, ou células gliais, contribui para as funções neurais, principalmente por meio do efeito isolante, de sustentação e de nutrição dos neurônios vizinhos. Se o sistema nervoso central fosse um biscoito com pedacinhos de chocolate e os neurônios fossem os pequenos pedaços de chocolate, a glia seria a massa do biscoito que preenche todos os espaços e assegura que os pedacinhos de chocolate se mantenham suspensos no seu local apropriado. De fato, o termo *glia* deriva da palavra grega para “cola”, dando a impressão de que a principal função dessas células é impedir que o encéfalo escape por nossos ouvidos” (BEAR *et al.*, 2017: 24).

⁶⁸ *Deleções* são mutações nas quais um trecho de DNA codificante é deletado ou perdido, resultando em proteínas disfuncionais. Diferentemente do que ocorre quando realizamos a leitura de uma mensagem lacunar construída com o código escrito de uma determinada língua, cujos elementos deletados podem ser recuperados inferencialmente dentro de certos limites (cf. *A casa est_vá pe_ando fogo qua_do os _ombeiros ch_garam*), a deleção de um único nucleotídeo de uma molécula de DNA pode desconfigurar o sentido da mensagem codificada em um gene, na medida em que o sistema de “leitura” das células – seu mecanismo de transcrição e tradução – tende a ignorar o elemento deletado, formando novos agrupamentos de três nucleotídeos a partir do ponto de deleção. Assim, uma sequência de bases do tipo (i) [...AAT GGC GTT CAC...], ao sofrer uma mutação por deleção de um de seus nucleotídeos [...AA_ GGC GTT CAC...], será lida pela maquinaria celular transcripcional como uma sequência do tipo (ii) [...AAG GCG TTC AC...], o que representa um *deslocamento de leitura* que resulta em um arranjo distinto de códons no RNAm transcrito e de aminoácidos no polipeptídeo traduzido (para uma explicação um pouco mais detalhada sobre o mecanismo de leitura do código genético, cf. a discussão da seção 2.1). Na síndrome de deleção do fragmento 22q13.3, a notação *22q13.3* se refere à localização cromossômica do trecho de DNA afetado pela deleção, que está situado no braço longo “q” (em

SHANK3 (também classificado como *ProSAP2*) afeta a síntese da proteína homônima *SHANK3* (cujas funções se relacionam com a organização do dispositivo de recepção de neurotransmissores das células pós-sinápticas e, conseqüentemente, com a eficiência da comunicação entre os neurônios [WILSON *et al.*, 2003]), fazendo com que o fluxo de informações em determinadas áreas do córtex cerebral não ocorra adequadamente (BONAGLIA, 2001). Como consequência, os indivíduos afetados apresentam graus variáveis de atraso global do desenvolvimento (incluindo alguns sintomas do espectro autista, cf. UCHINO & WAGA, 2013; MONTEIRO & FENG, 2017), manifestando muitas vezes um atraso severo no desenvolvimento da linguagem – e, em casos mais graves, a sua total ausência. Uma vez que atua na regulação da densidade de receptores pós-sinápticos cuja disfuncionalidade resulta em déficits linguísticos, o gene *SHANK3* pode ser considerado um representante do primeiro nível de ação gênica associada à linguagem. Um segundo nível do programa genético subjacente ao órgão da linguagem seria composto pelos genes que são ativados durante o processamento linguístico e que permitem regular e formatar o programa geral de desenvolvimento da linguagem, na medida em que são capazes de responder bioquimicamente aos estímulos ambientais, promovendo, através de seus produtos proteicos, uma organização mais detalhada e ajustada das redes neuronais responsáveis pela atividade linguística, de modo a construir uma gramática específica. Da tensão e da sintonia entre esses dois níveis de expressão gênica – o geral e o específico; o estrutural e o regulador; o ontogenético e o sensível ao ambiente –, emergiria um órgão da linguagem individualizado, cognitivamente dotado de uma língua-I.

oposição ao braço curto “p”) do cromossomo 22, na região 1, banda 3, sub-banda 3. Como modelo de leitura, cf. o exemplo da Fig. 4.1, referente à posição 2p25.3.

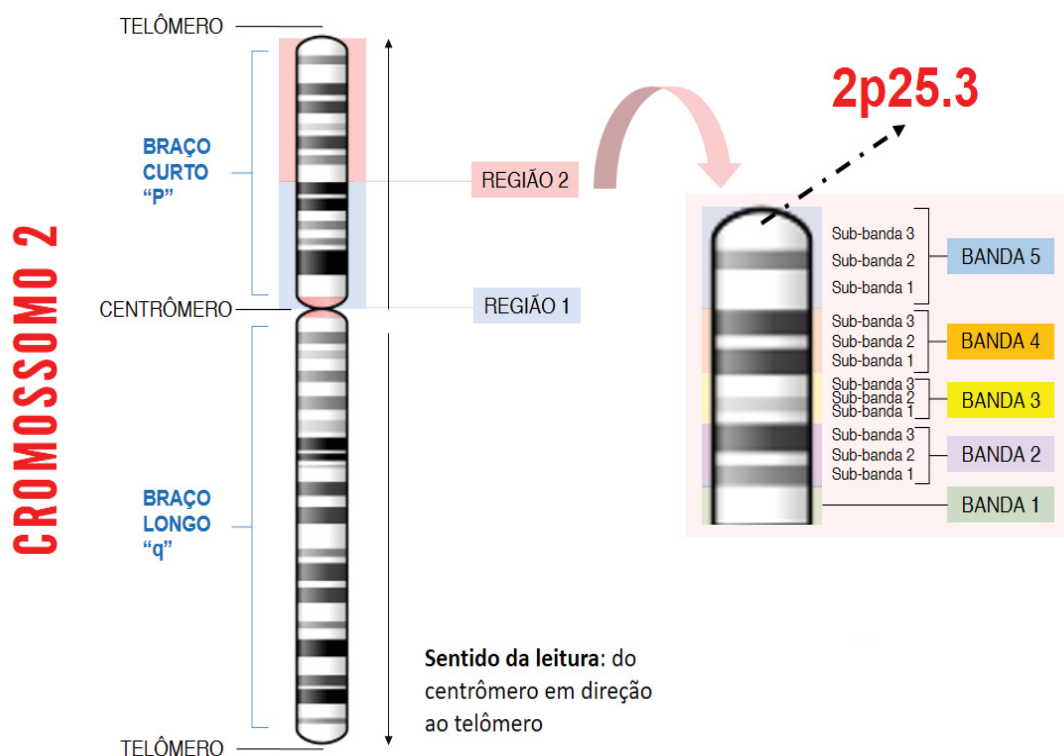


FIGURA 4.1. Notação cromossômica e exemplo de localização de um fragmento cromossômico (posição 2p25.3). (Fonte da imagem: <https://www.biomedicinapadrao.com.br/2020/02/como-ler-localizacao-de-um-genemutacao.html?m=1>)

O objetivo deste capítulo é apresentar alguns dos fatores genéticos que determinam as condições para o aparecimento e o desenvolvimento da linguagem no cérebro humano, o que também pode fornecer informações para a discussão de questões centrais a qualquer teoria que pretenda analisar a competência linguística a partir de uma perspectiva biológica, tais como a modularidade da mente/cérebro e as interações entre os diferentes componentes da arquitetura cognitiva. Como veremos na sequência, a caracterização desses genes parte tradicionalmente de estudos de análise clínica molecular de patologias nas quais a competência linguística aparece afetada, esteja esse déficit linguístico acompanhado de outros sintomas ou não (a deficiência de uma determinada proteína/enzima desregula as vias metabólicas das quais ela faz parte e normalmente afeta vários tecidos, de modo que distúrbios neurocognitivos que têm como causa mutações genéticas raramente possuem sintomas exclusivamente linguísticos).

Em alguns casos (como o do gene *FOXP2*, que será discutido em detalhes no cap. 5), padrões característicos de genealogia (i.e., de herança fenotípica) permitem que uma característica seja acompanhada ao longo das gerações de uma família e também podem

ajudar a traçar a relação existente entre um determinado transtorno linguístico (fenótipo) e o genótipo que o condiciona.

O estudo da etiologia genética de um determinado componente da cognição humana também pode partir da análise genômica comparativa interespecífica dos genes que a determinam – i.e., pode partir da caracterização e da análise de genes previamente identificados em outros organismos, cujas mutações dão origem a alterações neuronais e cognitivas semelhantes àsquelas encontradas em seres humanos (mutações essas que, dentro de certos limites bioéticos, podem ser manipuladas intencionalmente em organismos-modelo (como camundongos), a fim de se investigar as diretrizes moleculares de uma determinada característica neurocognitiva). No caso da linguagem, os programas de genômica comparativa têm como alvo aqueles genes que, em outras espécies, determinam a estruturação, o desenvolvimento, o funcionamento e a homeostase de núcleos cognitivos que de alguma forma são funcionalmente semelhantes aos centros responsáveis pelo processamento linguístico em humanos. O fato de o órgão da linguagem ser um dispositivo biológico exclusivo de nossa espécie não impede que encontremos semelhanças entre o programa genético que o condiciona e os programas genéticos que condicionam dispositivos cerebrais com funções cognitivas parecidas em outros seres vivos (em especial, aquelas relacionadas à atividade comunicativa) – afinal, a singularidade da faculdade da linguagem em humanos pode ter emergido de um *arranjo* exclusivo de genes, e não necessariamente de genes exclusivos (vide o caso do *FOXP2*, que, além de fazer parte da arquitetura genética humana, é encontrado em chimpanzés, camundongos e aves canoras, por exemplo). Até o momento, entretanto, o número de genes descritos cujas mutações parecem afetar o surgimento e o funcionamento do órgão da linguagem ainda é (relativamente) reduzido.

Como veremos logo a seguir, muitas disfunções linguísticas fazem parte de um conjunto maior de características que são afetadas pela mutação de um ou mais genes (erros inatos do metabolismo) e não é simples estipular o papel exato que um determinado gene exerce com relação à competência linguística em meio a um emaranhado de efeitos pleiotrópicos por ele condicionados. Se a mutação em um gene impede que os dendritos (Fig. 4.2) sejam formados corretamente, por exemplo, todo o sistema de transmissão e recepção de impulsos nervosos (potenciais de ação) dos neurônios pode ser prejudicado, afetando o desenvolvimento e o funcionamento das redes neuronais. Nesse caso, a eventual degradação dos núcleos responsáveis pela linguagem derivaria de uma má formação celular e o déficit linguístico emergiria apenas como mais um dos sintomas. Assim, a construção de um modelo

biolinguístico unificado e preciso que inclua todos os fatores genéticos que compõem os dois níveis de regulação gênica da linguagem acima referidos – o ontogenético-estrutural e o ambientalmente modulado – ainda está um pouco distante de ser efetivada. No entanto, o sequenciamento, a localização e a caracterização biomolecular dos genes identificados até agora, bem como a elucidação das funções neuroquímicas desempenhadas pelos produtos que codificam, ainda que sejam informações circunscritas a um grupo limitado de fatores genéticos e ainda que se refiram a genes que, em muitos casos, não se encontram ligados a funções exclusivamente linguísticas, representam avanços significativos para a compreensão do substrato neurofisiológico da linguagem e do programa genético que o condiciona – além de nos permitirem antever em que direção podem seguir as pesquisas futuras destinadas a investigar a natureza da linguagem a partir de suas bases genéticas.

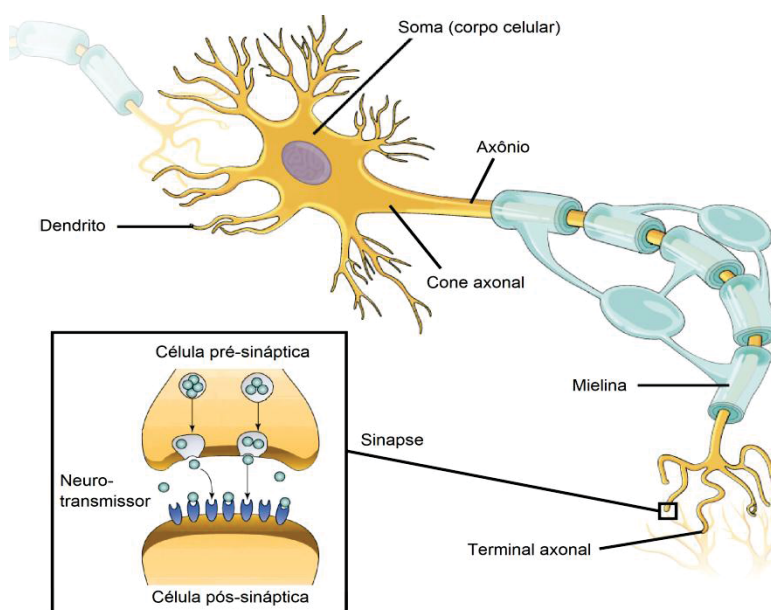


FIGURA 4.2. Estrutura geral do neurônio, com seus principais componentes. Em destaque, a região onde ocorre a sinapse. Os dendritos do neurônio pós-sináptico (receptor do impulso nervoso) recebem os neurotransmissores das terminações axonais da célula pré-sináptica (emissora do impulso). Caso o sinal excite o neurônio pós-sináptico, a informação é conduzida pelo axônio para que possa ser retransmitida a outro neurônio. (Fonte da imagem: <https://pt.khanacademy.org/science/biology/human-biology/neuron-nervous-system/a/overview-of-neuron-structure-and-function>. Acesso em 01/06/2021).

4.2. ALGUNS EXEMPLOS DE ERROS INATOS DO METABOLISMO QUE AFETAM A LINGUAGEM E OUTRAS FUNÇÕES COGNITIVAS

As subseções seguintes têm o objetivo de apresentar alguns genes que, ao sofrerem mutações, ocasionam problemas linguísticos (além de outros sintomas). Tentamos, na medida do possível, encontrar um equilíbrio entre a descrição do mecanismo genético subjacente e a

apresentação dos déficits linguísticos correspondentes. Para isso, lançamos mão de uma notação bioquímica um pouco mais detalhada em alguns pontos, mas sem nos aprofundarmos em pormenores excessivos. De todo modo, caso o(a) leitor(a) não se interesse por esse detalhamento mais fino do entrelaçamento entre determinados genes e determinados efeitos linguísticos, organizamos o texto de forma que é possível ir direto à seção 4.3 sem maiores prejuízos para a compreensão da discussão principal.

4.2.1. Gene *SLC17A5*

A doença de Salla é um transtorno metabólico hereditário progressivo que afeta o desenvolvimento de determinadas funções motoras e cognitivas – incluindo a linguagem – e surge como consequência de uma mutação do gene *SLC17A5*, localizado na região 6q14-q15 (VARHO *et al.*, 2002). Esse gene codifica a sialina, uma proteína transmembrânica cuja função é transportar o ácido siálico para fora dos lisossomos (pequenas vesículas intracelulares que carregam enzimas capazes de quebrar grandes biomoléculas em blocos de construção menores que o corpo pode reutilizar). A mutação em *SLC17A5* interfere no transporte do ácido siálico e faz com que ele se acumule nos lisossomos, impedindo que seja utilizado na formação da mielina (SONNINEN, 1999). A mielina forma a substância branca do cérebro e é um tecido com alto teor de lipídios que envolve e isola as fibras nervosas para que os impulsos sejam transmitidos na velocidade adequada (cf. Fig. 4.2). Na doença de Salla, o crescimento da mielina no cérebro é afetado e este processo neurodegenerativo pode levar à regressão do encéfalo (especialmente do cerebelo), provocando sintomas que começam a se manifestar após o primeiro ano de vida (desenvolvimento motor e cognitivo lentos na infância, levando a deficiências motoras e intelectuais na idade adulta).

Com relação especificamente à competência linguística, o desenvolvimento da fala se apresenta atrasado nas pessoas com doença de Salla e há uma progressiva degeneração dos centros de processamento da linguagem, que pode estar completamente ausente no indivíduo adulto. A capacidade de compreensão, no entanto, costuma ser menos afetada do que sua capacidade expressiva. Isso significa que as pessoas com a doença têm mais facilidade para entender os enunciados linguísticos do que para expressá-los verbalmente. As habilidades motoras orais são frequentemente afetadas e isso pode dificultar não apenas a articulação dos movimentos da fala, mas também a alimentação e o armazenamento de saliva. A perda da função cerebelar pode estar na base da degradação linguística e motora associada a essa

síndrome, uma vez que o cerebelo, além de ser um dos componentes responsáveis pelo controle motor, parece estar envolvido de alguma forma com funções cognitivas de natureza linguística, como percepção da fala e processamento semântico (MOBERGET & IVRY, 2016; DESMOND & FIEZ, 1998).

A doença de Salla é uma condição autossômica recessiva, o que significa que ambos os pais de uma criança afetada pela doença podem ser portadores saudáveis de um gene mutante (LANDAU *et al.*, 2004). Representando o alelo ativo (aquele que condiciona a produção adequada de sialina) por *S* e o alelo inativo (aquele que não sintetiza adequadamente sialina e, portanto, causa a doença de Salla) por *s*, temos as seguintes combinações genótípicas (com os fenótipos correspondentes entre parênteses): *SS* (fenótipo: indivíduo normal [não afetado]); *Ss* (indivíduo não afetado, portador do alelo mutante); *ss* (indivíduo afetado pela doença de Salla). Os indivíduos heterozigóticos *Ss* não são afetados pela doença de Salla porque o alelo que condiciona a produção normal de sialina (*S*) é dominante em relação ao alelo disfuncional (*s*) que condiciona a doença.

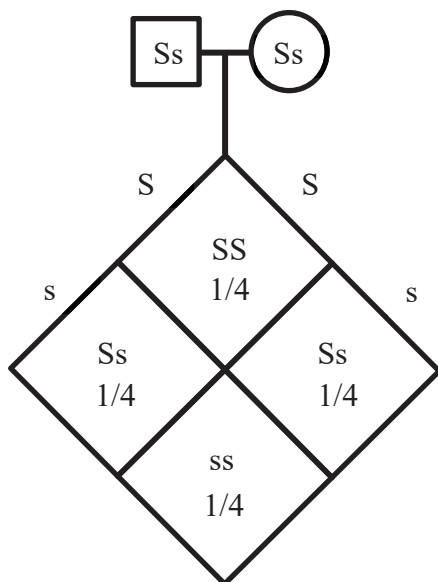


FIGURA 4.3. Probabilidade de descendência de um casal heterozigoto para o gene *SLC17A5*. O quadrado representa o pai; o círculo representa a mãe; os losangos representam a prole (de sexo indeterminado). *S* e *s* são os dois tipos de alelos transmitidos (através dos gametas). Dentro dos losangos, representam-se as possibilidades de fecundação e as probabilidades de combinação entre os alelos.

Proporção genotípica → 1 *SS* : 2 *Ss* : 1 *ss*

Proporção fenotípica → 3 não afetados : 1 afetado pela doença de Salla

4.2.2. Gene *PAH*

A fenilcetonúria (PKU) é uma doença hereditária recessiva causada por uma mutação do gene *PAH*, que está localizado na região cromossômica 12q24.1 e que é responsável pela

síntese da enzima fenilalanina-hidroxilase, cuja função é transformar a fenilalanina em tirosina (um precursor que serve de substrato para a formação de substâncias como a dopamina, a melanina e a adrenalina). A ausência da fenilalanina-hidroxilase nos indivíduos afetados pela mutação do gene *PAH* impede que a fenilalanina seja utilizada como substrato bioquímico para a formação de tirosina, provocando um desvio da via metabólica e fazendo com que substâncias deletérias se acumulem no sangue e no cérebro, causando lesões no sistema nervoso central (NUSSBAUM, 2016: 332-338). Caso não recebam tratamento adequado (focado essencialmente em uma dieta restritiva), fenilcetonúricos podem apresentar graus variados de distúrbios cognitivos que, no nível linguístico, afetam a capacidade de expressão, a memória verbal, a fluência fonêmica e diferentes processos semânticos (WHITE *et al.*, 2001). Por outro lado, indivíduos adequadamente tratados exibem processamento básico de linguagem intacto (acesso lexical, articulação fonêmica, construção de sentenças), mas podem apresentar déficits em habilidades pontuais de planejamento e raciocínio, sendo menos informativos na produção narrativa, mais lentos na compreensão metafórica e no cálculo do significado inferido e menos precisos na busca lexical focada (teste de Hayling⁶⁹), indicando que os déficits executivos em portadores de PKU não podem ser explicados por um acúmulo de déficits linguísticos de ordem inferior e que as funções executivas de raciocínio e planejamento envolvem recursos neurofisiológicos dedicados, em vez de simplesmente ser uma propriedade emergente de sistemas linguísticos de nível inferior (De FELICE *et al.*, 2018).

4.2.3. Genes *IDUA* e *SGSH*

As mucopolissacaridoses (MPSs) representam um conjunto de síndromes metabólicas causadas por mutações genéticas que levam à deficiência de determinadas enzimas

⁶⁹ O teste de Hayling consiste em um conjunto de duas tarefas em que o indivíduo testado precisa completar oralmente o conteúdo de sentenças que apresentam uma lacuna com relação à palavra final. Na primeira situação (*iniciação*), o indivíduo é solicitado a completar as construções com uma palavra adequada ao sentido da sentença, gerando uma medida de velocidade (de iniciação) de resposta. Na segunda situação (*supressão de resposta*), o indivíduo é solicitado a completar as frases com uma palavra que *não* esteja relacionada de forma alguma à sentença, resultando em construções semanticamente anômalas. A supressão de respostas adequadas gera uma medida de tempo de (inibição de) resposta e de velocidade de processamento linguístico (para mais detalhes, inclusive com exemplos de construções utilizadas nesse tipo de teste, cf. Burgess & Shallice [1996]).

lisossomais responsáveis pela metabolização adequada dos glicosaminoglicanos (GAGs), também chamados de mucopolissacarídeos (compostos bioquímicos pertencentes à classe dos açúcares). A deficiência enzimática faz com que as moléculas parcialmente degradadas de GAGs se acumulem nas células e no meio circundante, provocando alterações na adesão celular e na composição de alguns tecidos (que passam a ter uma consistência “viscosa”), causando problemas de natureza diversa. Nas mucopolissacaridoses do tipo I (MPS I) – uma classe de síndromes metabólicas que inclui a síndrome de Hurler (umas das formas mais graves de MPS, que se manifesta precocemente) e a síndrome de Scheie (um subtipo mais brando de MPS) –, uma mutação do gene *IDUA*, localizado na região cromossômica 4p16.3, provoca um déficit na síntese da enzima lisossomal alfa-L-iduronidase (IDUA), que é um dos elementos responsáveis pelo processamento metabólico dos glicosaminoglicanos (SCOTT *et al.*, 1995). A ausência de atividade enzimática acaba causando um acúmulo anormal de açúcares nos lisossomos, com consequências lesivas para o organismo.

O padrão de herança das MPS do tipo I é autossômico recessivo e os dois fenótipos (Scheie e Hurler) correspondem aos diferentes graus de deficiência enzimática. Na síndrome de Scheie, há uma ausência parcial de atividade enzimática e os indivíduos afetados desenvolvem uma forma menos agressiva da doença. Na síndrome de Hurler, a atividade metabólica da enzima IDUA está totalmente ausente, o que provoca a acumulação lisossomal de GAGs não degradadas, causando, entre outros sintomas, complicações cardiovasculares e respiratórias, deformações da estrutura musculoesquelética e surdez (NYHAN & HOFFMAN, 2020: 558-566). Além disso, a aglomeração de GAGs no tecido conjuntivo do cérebro causa alterações neurológicas que afetam o desenvolvimento psicomotor dos pacientes com síndrome de Hurler, fazendo com que eles apresentem uma deterioração cognitiva progressiva que os impede de adquirir uma competência linguística plena – uma degradação do sistema linguístico que está relacionada também com os déficits sensoriais (auditivos) e articulatórios que constituem a sintomatologia da síndrome (VAN DEN BROEK *et al.*, 2020).

Já a síndrome de Sanfilippo – ou mucopolissacaridose do tipo III (MPS III) – é uma doença autossômica recessiva causada pela ausência ou pela redução de atividade catalítica de uma das quatro enzimas responsáveis pelo metabolismo do heparan sulfato (YOGALINGAM & HOPWOOD, 2001) – um glicosaminoglicano de cadeia longa que, ao não ser degradado adequadamente nos lisossomos, acumula-se nas células, causando, entre outros sintomas, uma degeneração progressiva da atividade neurológica e das funções cognitivas associadas, com consequências potencialmente graves para a aquisição e o desenvolvimento da linguagem. Há

quatro tipos de MPS III (A, B, C e D), cada uma delas derivada do déficit de um tipo específico de enzima. Na síndrome de Sanfilippo tipo A (mucopolissacaridose IIIA), que é a variante mais comum e mais grave da doença, caracterizada por um colapso gradativo do sistema nervoso central, mutações no gene *SGSH*, localizado no trecho cromossômico 17q25.3 (SCOTT *et al.*, 1995), resultam na ausência ou na diminuição de atividade metabólica da enzima heparan sulfato sulfamidase (SGSH), uma das substâncias responsáveis pela “quebra” do heparan sulfato nos lisossomos. O acúmulo intracelular de longas cadeias de açúcar não totalmente degradadas afeta diretamente a homeostase das células, tornando-as disfuncionais. Os sintomas da síndrome de Sanfilippo tipo A normalmente começam a aparecer entre o primeiro e o sexto anos de vida – problemas comportamentais e atraso no desenvolvimento da fala, por exemplo (VALSTAR *et al.*, 2010) – e evoluem de forma progressiva, podendo levar a quadros severos de distúrbios cognitivos, tais como demência e deterioração da competência linguística.

4.2.4. Gene *DHCR7*

A síndrome de Smith-Lemli-Opitz (SLOS) é um distúrbio metabólico autossômico recessivo causado por uma mutação no gene *DHCR7*, que está localizado na região cromossômica 11q12-q13. O *DHCR7* é responsável pela codificação da 7-desidrocolesterol redutase, uma enzima utilizada em uma das etapas da via de biossíntese do colesterol (PORTER, 2008). O colesterol é um dos componentes estruturais das membranas celulares e faz parte da substância protetora que cobre e isola os axônios (mielina). Além disso, esse lipídio é utilizado como precursor de certos hormônios e ácidos digestivos e exerce um papel fundamental no desenvolvimento embrionário. A ausência de atividade enzimática causada pela mutação de *DHCR7* afeta a síntese adequada de colesterol e esse defeito no processo de metabolismo lipídico desregula as vias bioquímicas das quais o colesterol participa (Fig. 4.4), resultando em um espectro amplo de sintomas que caracterizam a SLOS: múltiplas anomalias congênitas, atraso cognitivo, transtornos comportamentais e distúrbios linguísticos de gravidade variável, relativos tanto à capacidade expressiva quanto à receptiva (De CLEMENTE *et al.*, 2013). Muitos dos sintomas característicos do distúrbio podem ser derivados do funcionamento irregular de algumas proteínas de sinalização responsáveis pelo controle do desenvolvimento embrionário, que só se tornam ativas funcionalmente se forem

modificadas pós-traducionalmente pela adição de um resíduo de colesterol (PORTER *et al.*, 1996).

A gradação dos traços fenotípicos associados à síndrome de Smith-Lemli-Opitz depende da natureza das mutações que atingem o gene DHCR7. *Mutações missense* – que alteram uma única base da sequência de DNA, fazendo com que a proteína sintetizada eventualmente possua um aminoácido diferente em uma determinada posição – correspondem à maioria dos casos de SLOS. Dependendo da localização da mutação e da importância do aminoácido substituído, mutações desse tipo podem afetar em maior ou menor grau a função da enzima e normalmente reduzem sua atividade metabólica (sem, entretanto, inativá-la totalmente). Já as *mutações nulas* – que impedem que a 7-desidrocolesterol redutase seja formada ou que resultam em produtos enzimáticos totalmente disfuncionais – são menos comuns e correspondem aos casos mais graves de SLOS, uma vez que a interrupção da síntese de colesterol afeta todas as vias da cadeia fisiológica da qual ele participa.

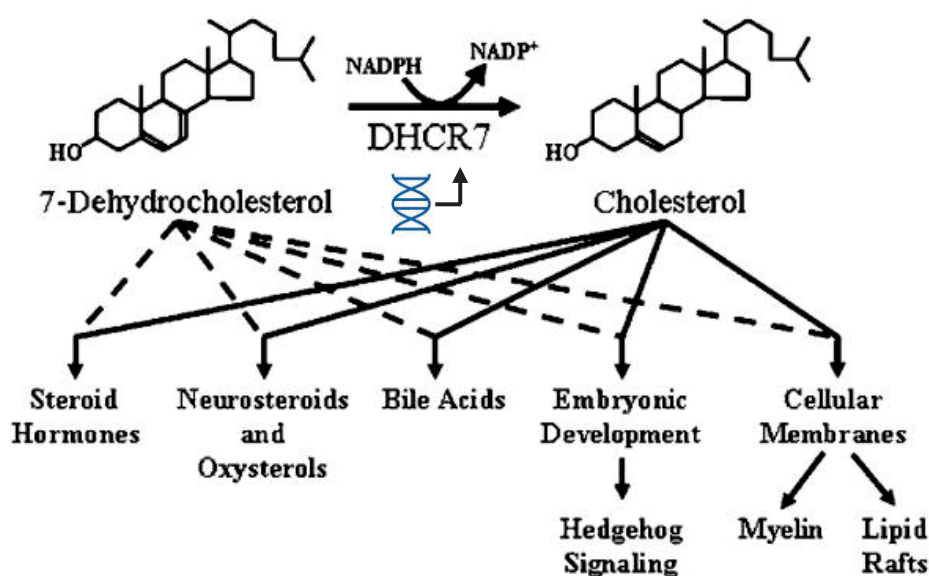


FIGURA 4.4. Via bioquímica da síndrome de Smith-Lemli-Opitz. Mutações no gene que codifica a enzima DHCR7 prejudicam a transformação de 7-desidrocolesterol (7DHC) em colesterol na última etapa de biossíntese desse lipídio. O colesterol exerce uma série de funções biológicas que podem ser afetadas pela sua carência, pelos efeitos tóxicos do excesso de 7DHC ou por uma combinação desses dois fatores (Imagem e texto adaptados de PORTER, 2008).

4.2.5. Genes *GALT*, *GALK* e *GALE*

A galactose é um carboidrato monossacarídeo de função energética que é obtido a partir da quebra (catabolismo) da lactose (o principal açúcar do leite): lactose (dissacarídeo) → glicose (monossacarídeo) + galactose (monossacarídeo). A galactosemia é uma condição autossômica recessiva caracterizada pela incapacidade enzimática do organismo de transformar quimicamente a galactose em glicose, provocando um acúmulo de galactose no sangue e fazendo com que a via de processamento natural desse açúcar sofra um desvio (cf. Fig. 4.5), dando origem a metabólitos tóxicos (como o galactitol) que podem levar o organismo a apresentar uma série de problemas, tais como hepatomegalia, catarata, atraso neurológico, comprometimento da atividade motora e atraso no desenvolvimento da linguagem (BOSCH, 2006; NYHAN & HOFFMAN, 2020: 425-434).

O metabolismo anormal da galactose é causado principalmente por mutações genéticas que provocam a deficiência ou o mau funcionamento de uma das enzimas responsáveis pelas etapas de transformação da galactose em glicose. Se tanto o pai quanto a mãe forem portadores assintomáticos de uma mutação genética causadora da galactosemia (i.e., se ambos forem heterozigotos (Gg) para um dos genes envolvidos na cadeia de processamento (via de Leloir) da galactose), cada filho do casal tem 25% de chance de ser homozigoto recessivo (gg) e de apresentar um dos fenótipos da doença. Existem três tipos principais de galactosemia, cada um deles definido pela ausência ou pela redução de atividade metabólica de uma determinada enzima. Na galactosemia clássica (tipo I), o acúmulo excessivo de galactose e de seus subprodutos é consequência da existência de uma mutação no gene *GALT*, localizado em 9p13, que codifica a enzima galactose-1-fosfato uridiltransferase (SHIH *et al.*, 1984). Sendo responsável pela forma mais grave da doença, esse erro metabólico causado pela mutação de *GALT* faz com que os indivíduos apresentem graus variáveis de comprometimento do controle motor, podendo levar à dispraxia da fala (NELSON *et al.*, 1991; WEBB *et al.*, 2003). Os outros dois tipos de galactosemia, menos agressivos, são decorrentes ou da deficiência da enzima galactoquinase (galactosemia tipo II), que é codificada pelo gene *GALK*, localizado na região cromossômica 17q24 (NOVELLI & REICHARDT, 2000), ou da ausência de atividade da enzima uridil-difosfo-(sigla: UDP)-galactose-4-epimerase (galactosemia tipo III), que é codificada pelo gene *GALE*, localizado na região cromossômica 1p36 (MACERATESI *et al.*, 1998). Mutações do gene *GALE* têm sido igualmente associadas a um déficit psicomotor e a um atraso geral no desenvolvimento da competência linguística e de outras funções cognitivas (ALANO *et al.*, 1998).

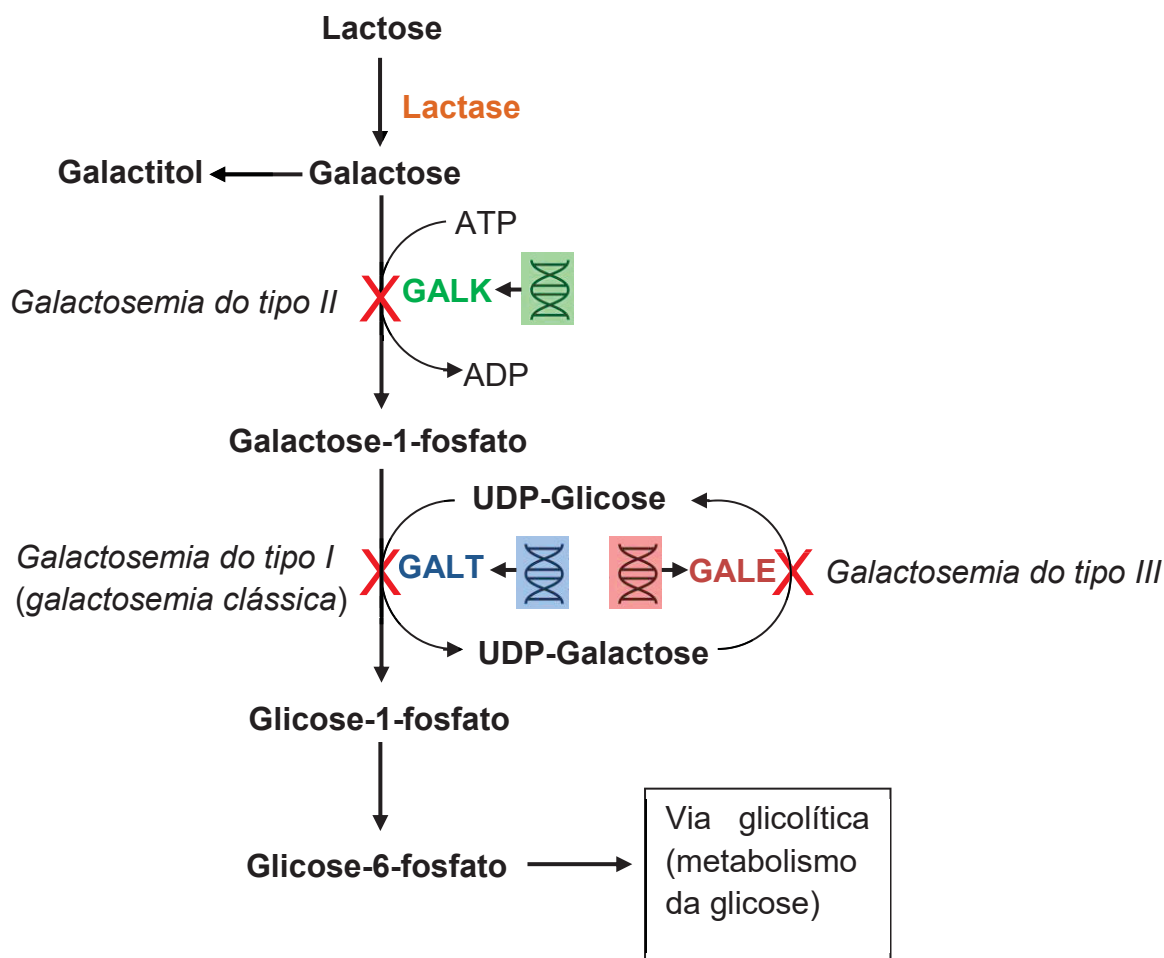


FIGURA 4.5. Via bioquímica do metabolismo da galactose (via de Leloir). Inicialmente, a galactose recebe um fosfato de um ATP (fosforilação) em uma reação catalisada pela galactose-1-fosfato uridiltransferase (GALK). Problemas nessa reação podem resultar no tipo II da galactosemia. A galactose-1-fosfato então reage com a UDP-glicose, formando glicose-1-fosfato e UDP-galactose. Essa reação é catalisada pela galactose-1-fosfato uridiltransferase (GALT) e a deficiência dessa enzima pode causar o tipo I da galactosemia, frequentemente associado a problemas de desenvolvimento da fala e dispraxia verbal. A glicose-1-fosfato produzida nessa etapa pode ser transformada em glicose-6-fosfato para então ser processada pela via glicolítica. Uma reação inversa ocorre quando a UDP-galactose é transformada em UDP-glicose (que pode então ser reutilizada na cadeia), em uma reação catalisada pela UDP-galactose-4-epimerase (GALE). Mutações no gene que codifica essa enzima podem acarretar o tipo III da galactosemia (Imagem e texto adaptados de TIMSON, 2016).

4.3. PAUSA PARA UM BALANÇO EPISTEMOLÓGICO E LINGUÍSTICO SOBRE O QUE TEMOS VISTO ATÉ AQUI

A explicitação dos mecanismos bioquímicos e das vias metabólicas que subjazem a determinadas anomalias genéticas caracterizadas por, entre outros traços fenotípicos, déficits linguísticos de natureza variada é um passo importante para a compreensão adequada da dimensão biológica da faculdade da linguagem. Os constantes avanços proporcionados pelas técnicas de investigação genômica fazem com que a engenharia microscópica da vida possa ser descrita de forma cada vez mais detalhada, fornecendo-nos um conjunto extenso de informações pormenorizadas sobre a estrutura molecular da matéria que compõe o mundo vivo. Porém, a relação de “causa e efeito” entre fatores genéticos e linguagem (gene X → propriedade linguística Y) ainda permanece em grande parte obscura em meio a este emaranhado de substâncias e reações químicas que sustentam e dão vazão aos fenômenos linguísticos. Assim, um dos desafios da (bio)linguística consiste em dar sentido aos dados obtidos pela genômica, explicando de que forma todo este arsenal de genes (do qual selecionamos apenas uma pequena amostra nas seções anteriores) se relaciona com a linguagem e qual é o papel específico que cada um deles exerce neste processo.

Atividades cognitivas como raciocinar, sonhar ou elaborar sentenças, além de se caracterizarem por operações mentais que se realizam na interface “virtual” do cérebro, correspondem a processos neurofisiológicos concretos executados pela interface “biofísica” do sistema nervoso. As propriedades fundamentais dessas duas interfaces – que, para os efeitos do que estamos discutindo aqui, estão sendo consideradas como instancias de uma coisa só (a esse respeito, cf. a nota 55, em que se discute brevemente a noção de *monismo fisicalista*) – derivam e são controladas pelo material genético. No entanto, como vimos nos exemplos das seções acima, genes codificam diretamente cadeias de aminoácidos (fenótipos primários concretos) e apenas indiretamente diretrizes cognitivas (fenótipos secundários abstratos). O problema fundamental apresentado à biolinguística, portanto, consiste em estabelecer a conexão etiológica entre os fenômenos imateriais da linguagem e os fatores materiais inscritos no DNA dos seres humanos, descobrindo a função abstrata (linguístico-gramatical) determinada por cada um desses genes.

Até aqui, temos discutido genes cujas mutações resultam em um quadro amplo de problemas que atingem os indivíduos afetados. A ausência de uma determinada substância (ou a presença de um produto gênico sintetizado de forma “equivocada”) gera uma espécie de

reação em cadeia que desestabiliza uma série de funções fisiológicas, incluindo aquelas relacionadas ao metabolismo cerebral e às atividades cognitivas. Especificamente com relação à cognição linguística, notamos que esses erros metabólicos inatos não atingem de modo exclusivo, sendo antes responsáveis por uma degradação mais ou menos generalizada do sistema nervoso. Nas mucopolissacaridoses, por exemplo, a redução ou o colapso da atividade enzimática provoca um acúmulo intra e extracelular de substâncias não catalisadas que acaba por afetar inúmeros componentes da arquitetura neurocognitiva, e não apenas aqueles relacionados ao processamento linguístico. Dado o espectro fenotípico amplo de características condicionadas direta ou indiretamente por esses genes, torna-se ainda mais complicado definir exatamente sobre qual componente da linguagem cada um deles está atuando. Além disso, a própria natureza dos déficits linguísticos que emergem dessas mutações – que, muitas vezes, dão origem a problemas generalizados de ordem fonêmica, morfossintática e semântica – dificulta ainda mais o estabelecimento de uma conexão precisa entre um determinado gene e a característica linguística (supostamente) condicionada por ele.

No entanto, se mudarmos o foco da pergunta que tem nos orientado até aqui – *por qual (ou quais) característica(s) da linguagem cada um desses fatores genéticos é (cor)responsável?* –, de modo a diminuir o grau de especificidade das informações que estão sendo solicitadas, talvez consigamos avançar sobre o assunto de forma mais segura e menos especulativa. Uma possível (re)formulação do nosso problema encontra-se em (1):

- (1) Qual é (ou quais são) o(s) efeito(s) linguístico(s) provocado(s) pela mutação de um determinado gene?

Note que agora não estamos nos comprometendo com uma relação causal imediata entre um determinado gene X e um determinado aspecto linguístico Y. Estamos simplesmente descrevendo quais são os efeitos linguísticos que emergem de uma determinada mutação genética. Se este gene atua diretamente sobre aspectos específicos da linguagem ou se está relacionado a aspectos neurocognitivos mais gerais – e de que forma tudo isso ocorre – são questões que precisam ser elucidadas caso a caso através de uma análise biológica e linguística aprofundada (algo que tentaremos fazer no próximo capítulo com o *FOXP2*).

A fim de oferecer um panorama – indubitavelmente incompleto, mas o mais atualizado possível – sobre o assunto, listamos, na Tabela 4.1 (*Erros inatos do metabolismo que afetam a linguagem e a cognição*), aqueles genes que, ao sofrerem mutação, dão origem a déficits linguísticos de natureza variada (entre outros sintomas cognitivos), com informações

sobre (i) localização cromossômica, (ii) efeitos linguísticos e cognitivos que derivam de suas mutações, (iii) denominação clínica do transtorno e (iv) referências bibliográficas relacionadas à sintomatologia cognitivo-linguística.⁷⁰

⁷⁰ Com algumas pequenas modificações (como a não especificação do fenótipo primário determinado pelos genes, ou seja, a não especificação das funções das proteínas sintetizadas), a Tabela 4.1 é uma versão adaptada da tabela presente em BURRACO, 2012: 240-247.

TABELA 4.1. Erros inatos do metabolismo que afetam a linguagem e a cognição (extraído e adaptado de BURRACO, 2012: 240-247)

Gene	Localização cromossômica	Efeitos linguísticos e cognitivos que emergem quando o gene sofre mutação	Denominação clínica do transtorno metabólico	Referências bibliográficas relacionadas aos aspectos linguísticos e cognitivos afetados
<i>THRB</i>	3p24.2	Alterações cognitivas e linguísticas; articularias, atenção e de aprendizagem; hiperatividade.	Síndrome de resistência ao hormônio tireoideano (RHT)	Mixon <i>et al.</i> (1992); Refetoff <i>et al.</i> (1993: 357-360)
<i>POMT1</i>	9q34.1	Atraso intelectual grave, acompanhado de problemas severos relacionados à aquisição da linguagem. Em uma das formas do distúrbio (distrofia muscular de Fukuyama), a criança pode apresentar desenvolvimento linguístico relativamente normal, mas ainda demonstra um atraso cognitivo significativo.	Síndrome de Walker-Warburg (um subtipo de distrofia muscular congênita)	Muntoni & Voit (2004)
<i>POMT2</i>	14q24.3	Idem	Síndrome de Walker-Warburg	Muntoni & Voit (2004)
<i>FCMD</i>	9q31	Idem	Síndrome de Walker-Warburg	Muntoni & Voit (2004)
<i>FKRP</i>	19q13.3	Idem	Distrofia muscular de Fukuyama	Muntoni & Voit (2004); Voit <i>et al.</i> (2004)
<i>GLUT1</i>	1p35-p31.3	Problemas leves de aprendizagem que, em alguns casos, podem afetar a aquisição de determinados componentes da fala	Encefalopatia por deficiência de <i>GLUT-1</i>	Wang <i>et al.</i> (2005)
<i>GALT</i>	9p13	Dispraxia verbal	Galactosemia clássica	Nelson <i>et al.</i> (1991); Webb <i>et al.</i> (2003); Timmers <i>et al.</i> (2011)

<i>GALK</i>	17q24	Déficit psicomotor; atraso no desenvolvimento da competência linguística e de outras funções cognitivas.	Galactosemia tipo II	Nyham & Hoffmann (2020): 425-434
<i>GALE</i>	1p36	Idem	Galactosemia tipo III	Alano <i>et al.</i> (1998) ; Nyham & Hoffmann (2020): 425-434
<i>SLC17A5</i>	6q14-q15	Atraso no desenvolvimento da fala; degeneração progressiva dos centros de processamento da linguagem, que pode estar completamente ausente no indivíduo adulto.	Doença de Salla	Adam & Wasserstein (2003); Moberget & Ivry (2016)
<i>ST3GAL5</i>	2p11.2	Epilepsia, perda auditiva e estagnação do desenvolvimento de áreas associadas à organização de competências psicomotoras, resultando, entre outros sintomas, na ausência da fala.	Deficiência de GM3 Synthase (Síndrome epiléptica infantil do tipo Amish)	Wang <i>et al.</i> (2016); Bowser <i>et al.</i> (2019)
<i>GALC</i>	14q31.3	Transtornos neurológicos que podem afetar o desenvolvimento e o funcionamento do sistema psicomotor.	Leucodistrofia de células globoides (Doença de Krabbe)	Orsini <i>et al.</i> (2000); Wenger <i>et al.</i> (2016)
<i>SLC6A8</i>	Xq.28	Atraso mental; problemas articulatórios e linguísticos significativos.	Síndrome de deficiência de creatina	Hahn <i>et al.</i> (2002); Salomons <i>et al.</i> (2003)
<i>SLC22A5</i>	5q31	Déficit intelectual; desordens cognitivas características do espectro autista; problemas articulatórios e linguísticos.	Deficiência primária de carnitina	Wang <i>et al.</i> (2001); Guevara-Campos <i>et al.</i> (2019)

<i>DHCR7</i>	11q12-q13	Atraso cognitivo, transtornos comportamentais e distúrbios linguísticos de gravidade variável, relativos tanto à capacidade expressiva quanto à receptiva.	Síndrome de Smith-Lemli-Opitz	Tierney <i>et al.</i> (2001); De Clemente <i>et al.</i> (2013)
<i>SUOX</i>	12q13.2	Forma clássica (de início precoce): déficit intelectual profundo; desenvolvimento mínimo da linguagem. Forma leve (de início tardio): desenvolvimento cognitivo e linguístico normais ou levemente afetados.	Deficiência isolada de sulfito oxidase (sulfocisteinúria)	Bindu <i>et al.</i> (2017); Tian <i>et al.</i> (2019)
<i>AUH</i>	9q22.31	Nos casos mais leves: atraso no desenvolvimento da linguagem e do sistema psicomotor. Nos casos mais graves: neurodegeneração progressiva associada a déficits linguísticos severos, ataxia e demência.	Acidúria 3-metilglutacônica tipo I	Gibson <i>et al.</i> (1998); Ijlst <i>et al.</i> (2002)
<i>GLDC</i>	9p24.1	Atraso intelectual, acompanhado de transtornos linguísticos (relacionados, principalmente, com a capacidade expressiva).	Hiperglicinemia não-cetótica (encefalopatia por glicina)	Van Hove <i>et al.</i> (2002); Hoover-Fong <i>et al.</i> (2004); Hennermann (2006)
<i>GCSH</i>	16q23.2	Idem	Idem	Idem
<i>AMT</i>	3p21.1-p21.2	Idem	Idem	Idem
<i>ASPA</i>	17p13.2	Neurodegeneração progressiva; atraso motor (dificuldade para falar, sentar, ficar em pé etc.); transtorno psicossocial e de aprendizagem. Em alguns casos, capacidade de se expressar usando palavras isoladas, mas incapacidade de se comunicar através de sentenças.	Doença de Canavan (degenerescência espongiiforme do sistema nervoso central)	Matalon <i>et al.</i> (1999); Bley <i>et al.</i> (2021)

<i>PAH</i>	12q24.1	Distúrbios cognitivos que, no nível linguístico, podem afetar a capacidade de expressão, a memória verbal, a fluência fonêmica e diferentes processos semânticos.	Fenilcetonúria	Azen <i>et al.</i> (1996); White <i>et al.</i> (2001); Janzen & Nguyen (2010)
<i>MCOLNI</i>	19p13.3-p13.2	Atraso global do desenvolvimento e/ou degeneração progressiva de habilidades cognitivas e motoras previamente adquiridas, com consequências severas para a emergência da linguagem ou para a capacidade expressivo-comunicativa.	Mucopolidose tipo IV	Altarescu <i>et al.</i> (2002); Jezela-Stanek <i>et al.</i> (2020)
<i>NAGA</i>	22q11	Crianças afetadas pelo tipo I da doença param de adquirir novas habilidades a partir de uma determinada fase (normalmente, antes de completarem 18 meses de idade) e passam a apresentar uma degradação progressiva das habilidades psicomotoras e cognitivas previamente adquiridas (<i>developmental regression</i>), o que provoca a interrupção da aquisição linguística e a deterioração dos componentes já adquiridos. O tipo II, que emerge normalmente na idade adulta, e o tipo III, uma forma infantil menos grave, associam-se a atrasos cognitivos e problemas psicomotores de grau leve ou moderado.	Doença de Schindler (Deficiência de alfa-N-acetilgalactosaminidase) tipos I (forma infantil grave), II (doença de Kanzaki) e III (forma intermediária)	Van Diggelen <i>et al.</i> (1987); Bakker <i>et al.</i> (2001)

<i>IDUA</i>	4p16.3	Atraso global do desenvolvimento cognitivo, que acaba dando lugar a um comprometimento intelectual e a uma aquisição incompleta da linguagem – uma aquisição deficitária que também deriva de problemas sensoriais (auditivos) e articulatorios.	Síndrome de Hurler; síndrome de Scheie (mucopolissacaridoses do tipo I)	Clarke (2002); Shapiro <i>et al.</i> (2018); Van Den Broek <i>et al.</i> (2020)
<i>SGSH</i>	17q25.3	Degeneração progressiva da atividade neurológica e das funções cognitivas associadas, com consequências graves para a aquisição e para o desenvolvimento da linguagem; alterações comportamentais significativas; demência.	Mucopolissacaridose do tipo IIIA (síndrome de Sanfilippo)	Van de Kamp <i>et al.</i> (1986); Valstar <i>et al.</i> (2010);
<i>MANBA</i>	4q22-q25	Atraso na emergência da linguagem, observando-se, em alguns casos, o surgimento de tiques fônicos (Síndrome de Tourette); deficiências auditivas; hiperatividade; déficit intelectual.	Beta-manosidose	Dorland <i>et al.</i> (1988); Sedel <i>et al.</i> (2006)

ARSA	22q13.33	Forma infantil: deterioração mental progressiva que começa a se manifestar antes da criança completar 30 meses, com consequências severas para a emergência da linguagem. Forma adulta (cujos sintomas começam a se manifestar após os 16 anos): instabilidade emocional; problemas de memória; distúrbios comportamentais e psiquiátricos (esquizofrenia). Normalmente, os indivíduos não apresentam déficits linguísticos.	Leucodistrofia metacromática	Kohn et al. (1988); Marcão et al. (2005); van Rappard et al. (2015)
GCDH	19p13.13	Macrocefalia; degradação neuronal dos gânglios basais e atrofia do lóbulo temporal. Hipotonia, distonia (as funções cognitivas se encontram menos afetadas do que as funções motoras); com relação à linguagem, a recepção se encontra menos prejudicada do que a expressão.	Acidúria glutárica tipo I	Kyllerman et al. (1994; 2004)
Até o momento, foram identificados mais de 70 genes relacionados à Síndrome de Leigh (cf. RAHMAN & THORBURN, 2015).	Localização cromossômica diversa, incluindo genes presentes em mitocôndrias.	Doença neurodegenerativa grave caracterizada por uma deterioração acelerada das funções motoras e cognitivas. As crianças portadoras normalmente morrem antes de completar 5 anos, mas há relatos de indivíduos mais velhos que, entre outros sintomas, podem apresentar declínio intelectual, déficit de memória, problemas de programação e expressão da fala (disfasia e disartria).	Síndrome de Leigh	Procaccio & Wallace (2004); Finsterer (2008); Baertling et al. (2014)

4.4. DISTÚRBIO(S) ESPECÍFICO(S) DE LINGUAGEM

Diferentemente do que ocorre com determinados tipos de erros inatos do metabolismo – cujos quadros fenotípicos são caracterizados por um espectro amplo de sinais clínicos (fisiológicos e cognitivos) inter-relacionados, e não apenas por déficits linguísticos de natureza variada –, o *distúrbio específico de linguagem* (DEL) ou *transtorno do desenvolvimento da linguagem* (TDL) é uma condição neurocognitiva nociva que afeta única e exclusivamente a aquisição e/ou o desenvolvimento da linguagem, sem que sejam observadas causas não linguísticas, tais como atraso global do desenvolvimento, autismo, problemas no desenvolvimento ou na execução de outras funções cognitivas (degeneração do sistema mnemônico ou déficits de atenção, por exemplo), problemas no desenvolvimento do sistema motor da fala ou do sistema receptor (perda auditiva/surdez), alterações neurossensoriais congênitas ou causadas por lesões neurológicas adquiridas etc. (COELHO *et al.*, 2013). Assim, diz-se que o DEL é *específico* exatamente pelo fato de emergir na ausência de outros déficits intelectuais, sensoriais ou motores e de quaisquer outras patologias, síndromes, danos cerebrais ou deficiências que possam prejudicar as condições neurofisiológicas e cognitivas ideais para a aquisição e o desenvolvimento normais da linguagem.⁷¹

O quadro de sinais clínicos que caracterizam o DEL corresponde a um conjunto extenso de déficits linguísticos – relacionados tanto à produção/expressão quanto à compreensão da linguagem – que pertencem a diferentes domínios (fonêmico, morfológico, sintático, semântico-discursivo) e que se manifestam em inúmeras configurações distintas, delimitando diferentes formas (ou recortes) do transtorno. Entre outros sintomas, os indivíduos afetados podem

[levar] um tempo maior no reconhecimento, recuperação, formulação e produção de palavras, devido à lentificação no processamento das informações, que pode estar relacionada a falhas nas representações semânticas e na organização cognitiva. Além disso, podem manifestar simplificações fonológicas, frequentemente desviantes (simplificações não observadas no processo normal de aquisição de linguagem); vocabulário restrito, com uso demasiado de dêiticos, perífrases e gestos representativos; estruturação gramatical simplificada e ordenação de palavras de forma não usual. Na compreensão, observam-se dificuldades em entender sentenças ou palavras específicas como marcadores espaciais ou temporais, realização de

⁷¹ Além disso, excluem-se do diagnóstico de DEL os transtornos linguísticos que derivam de condições socioeducativas inadequadas (como transtornos derivados de uma exposição insuficiente a estímulos linguísticos).

comandos linguísticos de forma incorreta, respostas incorretas sob questionamento e dificuldade em manter o tópico da conversação (CRESTANI *et al.*, 2013: 228-229).

É importante ressaltar que o distúrbio específico de linguagem pode se manifestar de inúmeras maneiras, circunscrevendo diferentes arranjos/combinações de déficits linguísticos. Alguns indivíduos afetados podem apresentar problemas de natureza fonêmica, por exemplo, enquanto outros apresentam problemas de natureza sintática ou semântica. O quadro clínico que caracteriza o DEL – seja com relação aos componentes da linguagem que são efetivamente afetados, seja com relação ao grau de comprometimento observado em cada componente (que pode ser leve, moderado ou grave) – é, portanto, variável. Por isso, dada a multiplicidade de manifestações clínicas que contempla, o termo DEL deve ser compreendido em sua heterogeneidade, uma vez que abarca uma gama variada de (subtipos) de distúrbios linguísticos. Alguns subtipos de DEL (com os seus respectivos sinais clínicos e experimentais) estão elencados na tabela a seguir, que se baseia nos critérios estabelecidos por Rapin & Allen (1983).

Distúrbios que afetam a capacidade expressiva

Dispraxia verbal

- Incapacidade maciça de fluência verbal
- Articulação severamente afetada, incluindo ausência completa de fala
- Compreensão normal ou muito próxima do normal

Distúrbio de Programação Fonológica

- Fala fluente, mas pouco inteligível
- Compreensão normal ou quase normal

Distúrbios que afetam a compreensão e a expressão

Déficit misto receptivo-expressivo ou distúrbio fonológico-sintático

- Fluência verbal perturbada
- Articulação da fala alterada
- Expressão limitada
- Sintaxe deficiente – frases curtas, omissão de palavras funcionais, agramatismo
- Compreensão deficiente em graus variados, mas com um grau de comprometimento menor do que o da expressão

Agnosia auditivo-verbal ou surdez verbal

- Compreensão da linguagem severamente afetada ou mesmo ausente
- Expressão limitada a frases curtas ou palavras isoladas, podendo estar totalmente ausente
- Articulação alterada
- Fluência verbal prejudicada

Distúrbios do processo central de tratamento e de formulação*Déficit semântico-pragmático*

- Fala fluente, muitas vezes logorreica⁷²
- Articulação normal
- Estrutura gramatical normal das frases
- Modos aberrantes de conversação
- Compreensão limitada de enunciados complexos

Déficit léxico-sintático

- Fala fluente, com ocasional pseudogagueira⁷³ devido a dificuldades de memória
- Articulação normal
- Jargão fluente⁷⁴ (em crianças pequenas)
- Sintaxe imatura, dificuldade com formulações complexas
- Compreensão limitada de sentenças complexas

TABELA 4.2. Classificação clínica dos diferentes subtipos de DEL (destacados em *itálico*) de acordo com Rapin & Allen (1983). Extraído e adaptado de CASTRO-REBOLLEDO *et al.*, 2004.

⁷² A *logorreia* (ou *logomania*) é um quadro compulsivo caracterizado por uma loquacidade excessiva e incessante. Essa “compulsão para a fala” muitas vezes acompanha quadros de confusão mental e é observada em determinados casos de surtos psicóticos.

⁷³ A *pseudogagueira* (ou *gagueira psicogênica*), diferentemente do que ocorre na gagueira comum, não é um distúrbio decorrente de problemas que atingem o sistema motor de produção da fala, mas sim um distúrbio que deriva de problemas cognitivo-psicológicos.

⁷⁴ O *jargão fluente* (*afasia de jargão* ou *jargonafasia*) é um tipo de afasia fluente em que a fala do indivíduo afetado, apesar de aparentemente continuar fazendo sentido para o emissor, torna-se incompreensível para os demais interlocutores. Pessoas que possuem essa condição podem realizar trocas de sílabas/sons no interior das palavras (substituição fonêmica) ou substituir integralmente uma palavra desejada por outra que soe como a original (ou que tenha alguma outra conexão – funcional, semântica etc. – com ela), podendo, também, substituí-la por conjuntos de sons aleatórios e sem sentido, resultando em neologismos (SWANBERG *et al.*, 2007).

Apesar das dificuldades relativas à conceptualização precisa do DEL – e considerando também toda a complexidade relativa à classificação de seus subtipos, que depende do estabelecimento de critérios adequados de identificação de características linguísticas pertinentes –, esta disrupção específica do sistema neurocognitivo responsável pela aquisição e pelo desenvolvimento da linguagem pode ser caracterizada de uma forma mais ou menos segura a partir de um ponto de vista clínico e experimental:

[...] uma criança apresenta um DEL se o transtorno não pode ser atribuído a nenhuma causa óbvia [uma lesão cerebral ou uma patologia metabólica, por exemplo], se só afeta alguma (ou algumas) de suas habilidades linguísticas, se suas execuções em tarefas relativas à linguagem são significativamente piores do que as que se referem a outras habilidades (principalmente cognitivas) e se os problemas linguísticos perduram ao longo do tempo, ainda que suas manifestações mudem de alguma forma (FRESNEDA & MENDOZA, 2005: S51-S52).

O fato de tal distúrbio afetar única e exclusivamente a linguagem tem uma importância fundamental para o problema que temos discutido até aqui, uma vez que seu caráter exclusivo/específico oferece condições ideais para que algumas das (possíveis) relações entre fatores genéticos e aspectos linguísticos sejam mapeadas de modo mais preciso. Apesar da significativa amplitude do quadro sintomatológico que caracteriza o DEL, todos os traços fenotípicos do transtorno dizem respeito apenas à linguagem, o que nos permite isolá-la dos demais domínios cognitivos durante a investigação de suas possíveis bases (onto)genéticas. A ideia básica – simples na concepção, mas de execução relativamente complexa, inclusive do ponto de vista tecnológico – consiste em identificar, a partir de problemas linguísticos sensíveis, erros genéticos (mutações) condicionantes, encontrando, assim, possíveis regiões do genoma humano responsáveis pela codificação de diretrizes linguísticas.

Entre as metodologias aplicadas no estudo dos condicionantes genéticos de DEL, destacam-se a *agregação familiar* e a *análise comparativa de gêmeos monozigóticos*.

A agregação familiar de uma determinada característica X – i.e., o modo como essa característica está distribuída entre os indivíduos de uma mesma família (em quais – e em quantos – indivíduos está presente ou não) – pode nos fornecer dados importantes sobre a existência ou não de fatores genéticos associados a X. Membros de uma mesma família compartilham parcelas significativas de material genético e, por isso, tendem a apresentar traços fenotípicos em comum. Se uma característica qualquer – como o DEL ou a tendência à depressão, por exemplo – possuir algum fator de natureza genética entre aqueles responsáveis por sua determinação (que pode ser multifatorial), espera-se que uma pessoa portadora de tal

característica possua um número maior de familiares afetados pela mesma condição do que uma pessoa não portadora (que, juntamente com seus familiares, formam o chamado *grupo controle* dos estudos de agregação familiar). Com relação especificamente ao DEL, o grau elevado de concentração familiar do distúrbio, observado frequentemente na literatura (cf. a Tabela 4.3), é um forte indício de sua natureza genética.

Investigadores	Fonte dos dados	Porcentagem de membros familiares afetados	
		Probando	Controle
Neils and Aram (1986)	Questionário	20	3
Tallal, Ross & Curtiss (1989)	Questionário	39	19
Tomblin (1989)	Questionário	23	3
Haynes & Naidoo (1991)	Entrevista	22	—
Tomblin & Buckwalter (1994)	Teste direto	21	—
Lahey & Edwards (1995)	Questionário	26	—
Plante, Shenkman & Clark (1996)	Questionário	38	0
	Teste direto	63	17
van der Lely & Stollwerck (1996)	Questionário	39	9
Rice, Haney & Wexler (1998)	Entrevista	22	7
Tallal <i>et al.</i> (2001)	Questionário	32	8
	Teste direto	30	7
Conti-Ramsden, Simkin & Pickles (2006)	Entrevista	35	—
	Teste direto	35	—
Barry, Yasin & Bishop (2007)	Questionário	32	6
	Teste direto	24	9
Lindgren, Folstein, Tomblin & Tager-Flusberg (2009)	Teste direto	52	—
Rice, Smith & Gayán (2009)	Teste direto	40	—

TABELA 4.3. Agregação familiar de DEL. Taxas de ocorrência de problemas linguísticos em famílias de crianças com DEL e em famílias de crianças sem DEL (grupo controle). Extraído de LEONARD, 2014.

A Tabela 4.3 relata o percentual de familiares com histórico de problemas de linguagem em 14 estudos diferentes. A maioria desses estudos, além de ter examinado a incidência de DEL nos membros da família da criança portadora do distúrbio (probando), examinou também a incidência de DEL em familiares de crianças não portadoras (grupo

controle). No entanto, também estão inclusos estudos que carecem de um grupo de comparação (na maioria dos casos porque estavam entre as primeiras pesquisas a empregar dados de testes diretos ou entrevistas). Como pode ser observado no conjunto dos estudos, os percentuais de familiares afetados foram significativamente maiores para as famílias do probando (+ DEL) do que para as famílias das crianças controle (– DEL), o que aponta para a existência de conexões entre DEL e fatores genéticos condicionantes – fatores estes que, possivelmente, estão sendo compartilhados pelos familiares afetados.

A porcentagem fornecida na tabela 4.3 para o estudo de Rice, Smith & Gayán (2009), por exemplo, é baseada nos resultados de uma pesquisa abrangente que incluiu testes de compreensão e de produção de linguagem. Esses pesquisadores também aplicaram testes com focos mais específicos e relataram a porcentagem de membros da família que poderiam ser considerados afetados a partir desses critérios. Para irmãos de até nove anos de idade, por exemplo, o percentual de indivíduos afetados foi de 72% para MLU, 71% para repetição de pseudopalavras, 57% para produção de marcadores de tempo gramatical e 46% para vocabulário receptivo, indicando graus elevados de agregação familiar com relação a diferentes subtipos de DEL (LEONARD, 2014). Já o estudo conduzido por Tallal *et al.* (2001) demonstrou que 13% das crianças DEL positivas investigadas eram filhos de pais não afetados, enquanto 40% tinham um dos pais afetados e 71,4% vinham de famílias cujos pais eram ambos portadores do distúrbio. Como veremos a seguir, as taxas de incidência mútua de DEL em gêmeos geneticamente idênticos (monozigóticos) são ainda mais acentuadas e endossam ainda mais o aspecto hereditário do transtorno.

Apesar das robustas evidências de agregação familiar de DEL, não são todos os indivíduos afetados que apresentam históricos familiares positivos para problemas de linguagem. Bishop & Edmundson (1986), por exemplo, descobriram que apenas 24% dos indivíduos “DEL positivos” investigados em seu estudo possuíam parentes de primeiro grau diagnosticados com algum tipo de distúrbio linguístico, enquanto 76% não possuíam parentes afetados e aparentemente eram os “primeiros” da família a apresentar o transtorno (LEONARD, 2014).

Como explicar de forma integrada as taxas de agregação familiar de DEL acima do normal e as porcentagens consideráveis de indivíduos DEL positivos que não possuem familiares afetados por problemas linguísticos? Uma explicação simples para essa aparente contradição consiste em assumir que os casos isolados de DEL derivam de mutações

genéticas “inéditas” na história familiar do indivíduo. Como vimos no início do Capítulo 2, as informações biológicas de uma determinada espécie são transmitidas de geração a geração através de mecanismos de replicação do material genético. Se uma determinada sequência desse material sofre alguma mutação durante a replicação, é provável que a característica por ela condicionada também seja modificada, afetando o indivíduo que a recebe “de presente” de seus genitores. Assim, não é de se estranhar que pais saudáveis tenham filhos portadores de doenças determinadas por condições genéticas deletérias; da mesma forma, não é surpreendente que um determinado tipo de DEL apareça “de repente” em uma família, uma vez que pode ser fruto de um cruzamento genético específico que deu origem a um indivíduo portador de mutações específicas em seu DNA.

Uma vez surgido, espera-se que esse fenótipo seja incorporado ao banco de características da família, dado que sua contraparte genética pode ser transmitida do indivíduo portador para os seus descendentes, aumentando, assim, a frequência do genótipo condicionante e a concentração familiar da característica condicionada. No entanto, os casos totalmente isolados de DEL – em que os indivíduos afetados não possuem nem ascendentes nem descendentes com problemas linguísticos – são relativamente comuns e representam um problema para a hipótese de uma etiologia exclusivamente genética para o DEL. Por isso, a hipótese mais aceita – e para a qual existe um conjunto mais robusto de evidências – é a de que alguns subtipos de DEL podem ser causados exclusivamente por fatores genéticos (FISHER *et al.*, 2003) enquanto outros podem ser causados por combinações de fatores genéticos e fatores ambientais/circunstanciais (condições do meio intrauterino; condições nutricionais da mãe durante a gestação; infecções neurológicas ou patologias cerebrais adquiridas pelo indivíduo afetado durante o período pré-natal ou durante a primeira infância etc.). De fato, a proposição de uma base multifatorial para o DEL, na qual cada fator – genético ou ambiental – exerce cumulativamente um pequeno efeito sobre o quadro final do distúrbio, parece responder mais adequadamente ao *continuum* fenotípico formado pelos diferentes tipos de DEL discutidos na literatura, uma vez que contempla um conjunto amplo de causas (hereditárias, congênitas, pré-natais, perinatais, pós-natais, endócrinas etc.). Além disso, tal abordagem multifatorial é a mesma adotada na análise de outros transtornos cognitivos genéticos, que raramente derivam de mutações genéticas isoladas, mas decorrem do efeito cumulativo de inúmeros fatores genéticos de risco modulados pelo ambiente (PLOMIN & RUTTER, 1998).

Os estudos envolvendo gêmeos monozigóticos (também chamados de gêmeos univitelinos, gêmeos idênticos ou gêmeos verdadeiros) são importantes para avaliar as participações relativas do genótipo e do ambiente na determinação do fenótipo. Como o próprio nome indica, os gêmeos monozigóticos se originam de um mesmo zigoto e, por isso, são geneticamente idênticos (Fig. 4.6). Assim, espera-se que uma característica condicionada por fatores genéticos seja expressa da mesma forma em gêmeos monozigóticos, a menos que sua expressão fenotípica seja influenciada também por fatores ambientais.

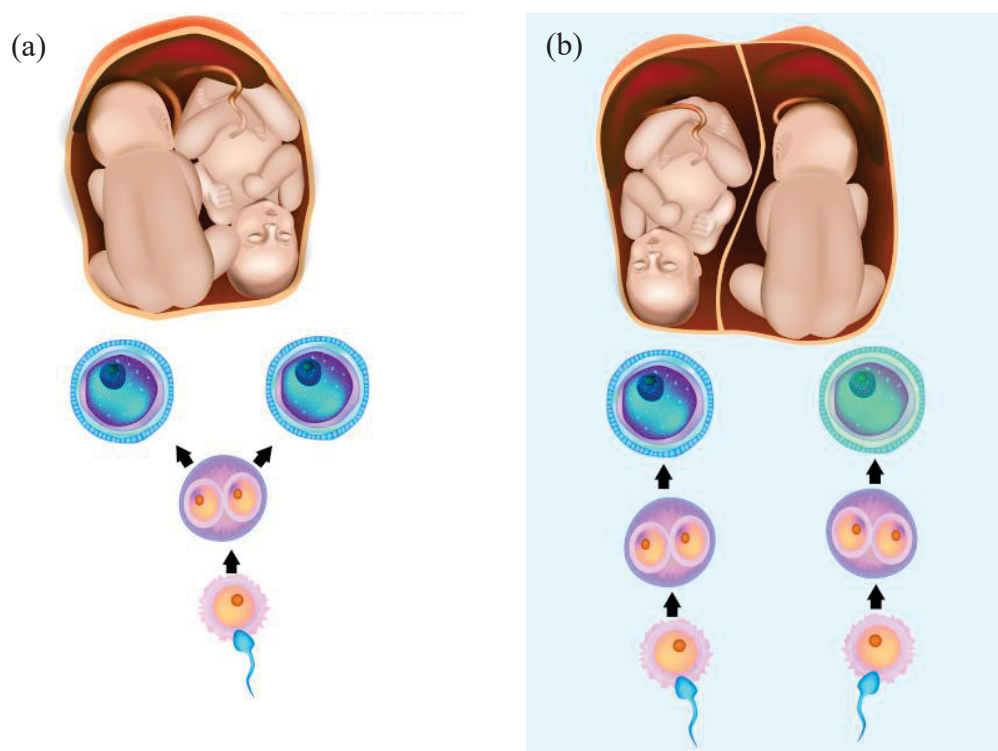


FIGURA 4.6. (a) Gêmeos idênticos (ou monozigóticos) são formados pela divisão precoce de um mesmo embrião e possuem conjuntos idênticos de genes. (b) Gêmeos fraternos (ou dizigóticos) são formados pelo desenvolvimento de zigotos diferentes (dois ou mais gametas femininos fecundados por diferentes gametas masculinos) e compartilham, em média, 50% de seu material genético. (Fonte da imagem: shutterstock.com. ID: 1608339985. Acesso em 11/10/2021).

A *taxa de concordância fenotípica entre gêmeos* expressa a porcentagem de pares de gêmeos que manifestam igualmente uma determinada característica. Quanto maior for a influência do genótipo na determinação da característica, maior deverá ser a taxa de concordância entre gêmeos monozigóticos – especialmente quando comparada à taxa de concordância observada entre gêmeos fraternos, que tende a ser menor já que os indivíduos são geneticamente distintos entre si. A taxa de concordância referente ao tipo sanguíneo do

sistema ABO, por exemplo, é de 100% entre gêmeos idênticos e de aproximadamente 65% entre gêmeos fraternos, indicando que o tipo de sangue (A, B, AB ou O) é uma característica fortemente determinada por fatores hereditários⁷⁵. Já o uso predominante da mão direita ou da mão esquerda, ao que tudo indica, recebe forte influência do ambiente compartilhado, uma vez que as taxas de concordância entre gêmeos idênticos (80,49%) e entre gêmeos fraternos (79,27%) são essencialmente as mesmas (PFEIFER *et al.*, 2022).

Com relação especificamente ao DEL, os dados provenientes dos estudos comparativos entre gêmeos monozigóticos (MZ) e dizigóticos (DZ) fornecem evidências suficientemente robustas para supor que os fatores genéticos exercem um papel importante na determinação de algumas classes de comprometimentos linguísticos. A maioria dos estudos de concordância citados na sequência – retirados de Leonard (2014) – partiu da análise contrastiva entre um gêmeo que era afetado por um distúrbio de linguagem (definido na maioria dos casos como apresentando DEL) e o gêmeo correspondente, que foi avaliado para determinar a concordância ou não do traço fenotípico sob análise.

Em um estudo com 57 pares de gêmeos, Lewis & Thompson (1992) encontraram taxas de concordância de 86% para gêmeos MZ e 48% para gêmeos (DZ) com relação ao DEL. Gêmeos MZ afetados por problemas de linguagem também eram mais propensos a apresentar os mesmos tipos de dificuldades não linguísticas (como as relativas à aprendizagem escolar). Bishop *et al.* (1995) analisaram 90 pares de gêmeos e encontraram taxas de concordância de 70% para pares MZ e de 46% para pares DZ, com forte evidência de herança de fatores genéticos aplicando-se a crianças com distúrbios fonêmicos e déficits na capacidade expressiva. No entanto, problemas referentes à percepção/compreensão da linguagem não pareciam atender ao critério de herança genética. Tomblin & Buckwalter (1998), em um estudo envolvendo 62 pares de gêmeos e três conjuntos de trigêmeos, relataram uma taxa de concordância de 96% para gêmeos MZ e 69% para gêmeos DZ, confirmando, assim, as conclusões de estudos anteriores quanto à existência de fortes determinantes genéticos ligados ao DEL. Em um estudo com gêmeos que incluiu 165 crianças com problemas de expressão, de recepção ou distúrbios dos sons da fala, DeThorne *et al.* (2006) encontraram maiores taxas de concordância entre gêmeos MZ do que entre gêmeos DZ para cada um desses subtipos de distúrbios – com destaque para as elevadas taxas de concordância monozigótica detectadas

⁷⁵ De fato, os diferentes grupos sanguíneos do sistema ABO são determinados geneticamente por três tipos de alelos que podem ocupar alternativamente o mesmo *locus* gênico e que são responsáveis pela síntese (ou não) de diferentes proteínas (aglutinogênios) presentes nas membranas da hemácias.

com relação aos distúrbios de linguagem expressiva (89%) e aos distúrbios dos sons da fala (86%), endossando, mais uma vez, as hipóteses relativas à etiologia genética do DEL.

Como observado em inúmeros estudos, crianças com DEL muitas vezes possuem limitações graves relacionadas à memória fonológica e à competência gramatical (computação). Bishop *et al.* (2006) realizaram um amplo estudo de concordância com 173 pares de gêmeos para determinar se esses dois tipos de comprometimento, como também alguns problemas de vocabulário, atendiam ou não aos critérios de herança genética. As competências que mostraram herdabilidade clara foram a repetição de pseudopalavras (uma habilidade que normalmente é associada à capacidade da memória fonológica), a produção de flexões e a compreensão da estrutura da frase (duas capacidades ligadas à computação gramatical). As taxas comparativas não mostraram praticamente nenhum efeito do ambiente na determinação dessas características. A análise contrastiva dos dados, no entanto, revelou que, embora tanto a memória fonológica quanto a computação gramatical apresentassem evidências de hereditariedade, elas eram separáveis e, até certo ponto, autônomas. Algumas crianças apresentavam déficits em ambas as áreas, mas outras apresentavam déficits isolados em uma competência, sem que fosse encontrado um déficit concomitante na outra. Diferentemente dos resultados encontrados para os déficits de memória fonológica e de computação gramatical, que apontam para a existência de fatores genéticos envolvidos na determinação de ambas as competências, os problemas de vocabulário mostraram um padrão inverso, com influência significativa do ambiente compartilhado e, aparentemente, sem influência genética.

Crianças com DEL frequentemente exibem déficits motores mais ou menos sutis, apresentando problemas de coordenação motora. A fim de determinar se tais limitações motoras possuíam uma base genética com fatores que podiam corresponder aos observados nas limitações referentes à repetição de pseudopalavras, Bishop (2002a), utilizando uma metodologia de concordância dupla, descobriu que baixos desempenhos em tarefas de movimentação de pinos e de repetição de pseudopalavras mostravam taxas de concordância e de variação gênica semelhantes entre gêmeos, mostrando que poderia haver um componente genético comum subjacente a algumas formas de DEL e determinados problemas de coordenação motora.

Apesar da dificuldade em se classificar os diferentes tipos de DEL, os estudos de concordância entre gêmeos permitiram a identificação mais precisa de subgrupos de crianças

afetadas por distúrbios da linguagem. Em estudos tradicionais de subgrupos de DEL, por exemplo, a constatação de que algumas crianças apresentam simultaneamente problemas de memória fonológica e de computação gramatical enquanto outras apresentam apenas um desses problemas pode servir como obstáculo para o estabelecimento de subgrupos bem definidos, na medida em que não é possível estabelecer até que ponto os déficits estão relacionados ou não. No entanto, os achados de concordância entre gêmeos, como os observados no estudo de Bishop *et al.* (2006), indicam que esses dois tipos de déficits podem ser vistos como distintos – cada um deles determinado por um conjunto específico de fatores genéticos e com o seu próprio padrão de herança –, não obstante o fato de que alguns indivíduos podem acumular os fatores genéticos determinantes e apresentar ambos os tipos de problema. Outra contribuição dos estudos de concordância entre gêmeos monozigóticos e dizigóticos é a percepção de que uma deficiência em uma determinada competência linguística pode estar mais relacionada geneticamente a um problema de ordem não linguística do que a um problema de ordem linguística. O efeito pleiotrópico de um fator genético – i.e., o conjunto de características fenotípicas que são determinadas por ele – pode ser bastante amplo, derivando traços com aparentemente pouca relação entre si. A análise de Bishop (2002a), por exemplo, conseguiu detectar que existe uma aparente base genética compartilhada entre a capacidade de repetição de pseudopalavras (tributária da memória fonológica) e determinada habilidade motora, sem que um compartilhamento semelhante tenha sido observado entre a repetição de pseudopalavras e a computação gramatical, que parecem estar relacionadas a competências linguísticas geneticamente independentes (LEONARD, 2014).

Como dito anteriormente, alguns tipos de DEL parecem derivar diretamente de um número reduzido de genes mutados (eventualmente, apenas de um, como no caso do quadro de déficits linguísticos causado por mutações no *FOXP2*⁷⁶), enquanto outros parecem emergir da combinação e do efeito cumulativo de fatores genéticos influenciados por fatores

⁷⁶ Apesar de ser caracterizado por uma série de comprometimentos linguísticos, o quadro fenotípico relacionado a problemas no *FOXP2* normalmente não é definido como um distúrbio específico de linguagem no sentido estrito do termo, uma vez que os indivíduos afetados apresentam problemas motores e orofaciais importantes, além de outras limitações que se estendem para além das descrições atuais de indivíduos com DEL. Atualmente, o distúrbio condicionado por mutações no *FOXP2* é melhor caracterizado como uma dispraxia orofacial. Além disso, alguns trabalhos de genética molecular em indivíduos que apresentavam os sintomas mais clássicos de DEL não encontraram evidências de envolvimento do *FOXP2* no distúrbio (O'BRIEN *et al.*, 2003).

ambientais. Neste último caso, cada gene (ou fator genético) contribui individualmente para a caracterização do quadro de DEL, que adquire um padrão de herança quantitativa (acumulativa), caracterizado por um *continuum* de fenótipos (Fig. 4.7). Assim, o DEL emerge como o resultado de um ou vários déficits diferentes (BISHOP, 2002b) – auditivos, fonológicos ou morfossintáticos –, a depender da combinação e da quantidade de mutações genéticas específicas acumuladas e das condições ambientais.

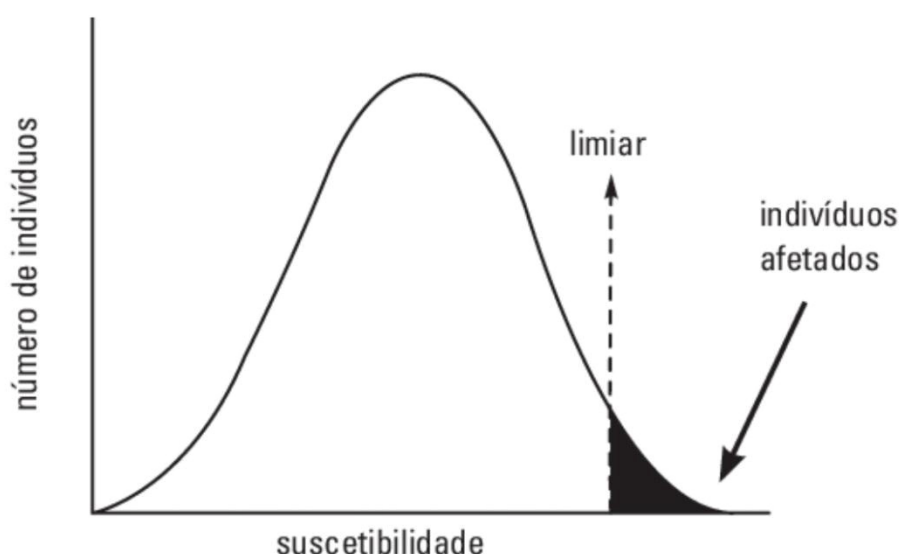


FIGURA 4.7. Modelo de limiar para herança multifatorial de DEL. Segundo este modelo, existe um limiar de fatores de suscetibilidade – genéticos e ambientais – que, quando ultrapassado, determina a manifestação do transtorno. Quanto maior for o número de fatores de risco acumulados em um indivíduo, maiores serão as chances de ele ser afetado pelo distúrbio. Abaixo do limiar da curva de distribuição normal, espera-se que o indivíduo tenha fenótipo normal.

Essa “imprecisão” fenotípica – caracterizada por um quadro de sintomas variáveis que se manifestam em diferentes graus de gravidade – sempre foi um obstáculo para a identificação das causas genéticas do DEL (afinal, para que se encontrem as causas biomoleculares de um problema genético é preciso que esse problema esteja bem definido quanto ao fenótipo que determina na superfície fenotípica). O déficit que normalmente é postulado como a deficiência central do transtorno, no entanto, é aquele que afeta a memória fonológica de curto prazo (MONACO, 2007; NEWBURY *et al.*, 2009) – uma competência que, entre outras funções, determina a taxa de aquisição lexical e a taxa de aquisição sintática (BADDELEY *et al.*, 1998). Além disso, problemas nesta modalidade de memória fonológica parecem estar intimamente relacionados com o aparecimento de outros distúrbios cognitivos

que afetam áreas ligadas à linguagem, como a dislexia. Assim como outras competências cognitivas, fortes evidências têm sido apresentadas a favor da existência de uma contraparte genética responsável pela memória fonológica de curto prazo (BISHOP *et al.*, 1999). Um dos achados mais importantes relacionados a essa competência adveio das investigações conduzidas pelo SLI⁷⁷ Consortium, que realizou (em 2002 e 2004) uma série de investigações de detecção de genes associados ao DEL.

O objetivo principal do SLI Consortium (2002; 2004) consistia em mapear genomicamente traços quantitativos (genes ou *loci* gênicos) ligados a determinadas manifestações fenotípicas do DEL, relacionadas principalmente às capacidades linguísticas de recepção e de expressão e à capacidade de repetição de pseudopalavras (habilidades que, não raramente, encontram-se degradadas nos indivíduos portadores do distúrbio). Normalmente, os estudos preliminares de mapeamento genético utilizam ferramentas matemáticas (frequência de recombinação etc.) para calcular o “endereço” cromossômico aproximado do(s) fator(es) genético(s) que condiciona(m) um determinado fenótipo. Apesar de extremamente útil, uma vez que permite localizar importantes regiões gênicas relacionadas a fenótipos específicos, delimitando áreas de interesse para a pesquisa biomolecular, esse mapeamento inicial acaba levando à identificação de segmentos que correspondem a intervalos cromossômicos relativamente extensos, que requerem estudos de acompanhamento de maior resolução e precisão (NEWBURY & MONACO, 2010). A ideia central de estudos como o do SLI Consortium consiste em realizar varreduras em regiões cromossômicas específicas de indivíduos que apresentam o mesmo tipo de déficit linguístico e compará-las com as sequências de DNA das regiões cromossômicas correspondentes de indivíduos não afetados, a fim de encontrar mutações significativas nestes segmentos.

Os resultados mais importantes das investigações promovidas pelo SLI Consortium dizem respeito a dois *loci* que foram identificados nos cromossomos 16 e 19. O locus mais significativo com relação ao DEL seria a região 16q24 (associada ao subtipo DEL₁ do distúrbio), que estaria preferencialmente ligada ao fenótipo de repetição de termos não linguísticos – uma habilidade que é utilizada para avaliar o funcionamento da memória fonológica (que, como mencionamos, está no núcleo de diversas manifestações de DEL). Essa ligação entre mutações na região 16q24 e problemas de memória fonológica de curto prazo – mais especificamente, problemas em tarefas de repetição de pseudopalavras – também foi

⁷⁷ SLI (do Inglês, *Specific Language Impairment*): distúrbio específico de linguagem (DEL).

identificada em estudos posteriores (como os de Monaco, 2007 e Newbury *et al.*, 2009), que promoveram análises ainda mais detalhadas no locus, chegando a identificar dois genes candidatos, *ATP2C2* e *CMIP* (NEWBURY, 2009). Além disso, foram detectadas pelo SLI Consortium ligações significativas de três variáveis da competência leitora (leitura básica, ortografia e compreensão leitora) com o locus mais amplo 16q – ligações essas que também foram identificadas no estudo de Monaco (2007). Esses achados são particularmente importantes devido à alta frequência de problemas de leitura e ortográficos em crianças com DEL, indicando que talvez existam segmentos cromossômicos (e eventualmente genes) ligados simultaneamente a déficits de fala e de leitura/escrita. O locus 19q13 (associado ao fenótipo DEL2) parece ser menos estatisticamente relevante e suas variáveis indicam uma ligação com o baixo desempenho em tarefas de capacidade expressiva (SLI Consortium, 2002) e, talvez, uma ligação com a capacidade de repetição de pseudopalavras (SLI Consortium, 2004), enquanto que a relevância estatística de sua relação com as demais variáveis fenotípicas analisadas seria menor. Estudos posteriores (como o de Falcato *et al.*, 2008) trouxeram ainda mais evidências de ligação entre o locus 19q13 e a capacidade expressiva. Nenhum desses *loci* – 16q24 e 19q13 – estaria ligado ao quociente de inteligência não verbal, excluindo-se, assim, deficiências intelectuais de base.

Transtornos em determinadas funções linguísticas podem ser utilizados como guias de exploração do código genético, servindo para os pesquisadores como uma tela dupla em que, de um lado, são vistos os problemas de superfície (fenotípicos), enquanto do outro se buscam erros de programação em determinadas áreas do DNA. No entanto, apesar das evidências quanto à existência de fatores genéticos envolvidos diretamente na determinação de alguns tipos de DEL, a variabilidade de seu quadro de sintomas e a existência de indivíduos que aparentemente não apresentam mutações genéticas identificáveis são fortes indícios da natureza multifatorial do transtorno. Mesmo quando as funções bioquímicas dos genes candidatos (transporte de proteínas, ativação ou inativação de outros genes, remoção de elementos químicos intracelulares etc.) forem mais bem caracterizadas do ponto de vista molecular, explicar como distúrbios nessas funções primárias levam a alterações em comportamentos complexos de linguagem será um imenso desafio, uma vez que fatores extragenéticos podem influir de modo significativo no quadro final dos transtornos. Entretanto, o desenvolvimento de modelos teóricos de efeitos epigenéticos e de regulação gênica pode ajudar a resolver esse desafio multifatorial.

Modelos epigenéticos levam em consideração as mudanças no padrão de expressão gênica que não decorrem de modificações na sequência de bases do DNA (mutações), mas que são moduladas por fatores ambientais. Durante o desenvolvimento, o DNA que compõe os genes acumula “marcas” químicas que modulam as taxas de produção gênica, determinando em que ocasiões e em que quantidade os genes serão expressos, como resposta a alterações ambientais que solicitam determinadas demandas de produtos gênicos. Fatores ambientais podem, por exemplo, desencadear a adição de grupos metil ($-\text{CH}_3$) a um determinado segmento de DNA (metilação), provocando a sua compactação e diminuindo a sua expressão naquele segmento (a maquinaria de transcrição só é capaz de ler sequências de DNA que estejam descompactadas). Assim, o conjunto de experiências acumuladas pelo indivíduo – principalmente na vida intrauterina e na infância – modula e reorganiza seu epigenoma, tornando-o altamente especializado e fazendo com que mesmo gêmeos idênticos, que possuem o mesmo DNA de base, mas que crescem sob um conjunto relativamente único de experiências ambientais moduladoras, possuam taxas distintas de produtos gênicos sintetizados, o que resulta em diferenças fisiológicas, cognitivas e comportamentais.

Rice (2012) propôs um desses modelos epigenéticos para o DEL, denominado *Modelo de interrupção (ou interrupção) da sinalização de crescimento* (“*Growth Signaling Disruption*”). Nesse modelo, o desenvolvimento linguístico requer (i) um mecanismo de inicialização, responsável pela aquisição e pelo surgimento da linguagem, (ii) uma função de aceleração, que controla o ritmo de desenvolvimento das diferentes fases da aquisição linguística e (iii) uma função de desaceleração, que é ativada na pré-adolescência, quando a janela de aquisição normalmente se fecha. Com base em sua análise comparativa de diferenças de trajetória de crescimento entre crianças com DEL e crianças com desenvolvimento típico, Rice propôs que a ativação do mecanismo de inicialização é atrasada nas crianças afetadas devido a algum fator que interrompe a sinalização para o crescimento. A função de aceleração parece operar como em crianças típicas, mas não há plasticidade neurocognitiva suficiente, resultando em efeitos de teto prematuros quando a desaceleração começa, o que provoca a estabilização dos déficits linguísticos. Quais fatores ambientais funcionam como gatilhos epigenéticos determinantes do DEL e como essas taxas diferenciais de expressão gênica levam a uma assincronia nos mecanismos de aquisição da linguagem ainda são questões sem resolução. Contudo, o modelo de Rice oferece uma matriz teórica importante, pois não é guiado apenas pelo comportamento linguístico das crianças com DEL, mas também por uma tentativa de elucidar como esse comportamento muda ao longo do

tempo em função de fatores epigenéticos de regulação do DNA, o que pode explicar mais eficientemente os casos de DEL que não decorrem de modificações visíveis nos genes (LEONARD, 2014). Estendendo um pouco essa análise, é possível prever que a ausência de um *input* linguístico mínimo poderia influenciar negativamente a aquisição da linguagem desde suas bases genéticas, uma vez que representaria uma ausência quantitativa e qualitativa de estímulos ambientais capazes de acionar a atividade de certos genes a fim de modular adequadamente taxas diferenciais de expressão gênica cerebral de acordo com as exigências de cada fase do desenvolvimento infantil.

Por fim, assim como fizemos para os casos de erros metabólicos que afetam de alguma forma a linguagem, apresentamos na tabela 4.4 uma amostra de genes e *loci* gênicos que possuem fortes ligações com distúrbios especificamente linguísticos⁷⁸. Além disso, também elencamos um conjunto de genes relacionados à dislexia, devido à imbricação do transtorno com determinadas áreas da linguagem. Trata-se, obviamente, de um panorama geral e não exaustivo das possíveis relações entre genes e déficits linguísticos. Hoje, o conjunto de informações sobre esse tema/problema conta com um extenso banco de dados a ser explorado em estudos futuros. O *Human Phenotype Ontology* (<https://hpo.jax.org/app/>) é uma ferramenta de pesquisa *on-line* que permite realizar buscas cruzadas entre genes, doenças e características fenotípicas (sintomas clínicos). O *link* a seguir dá acesso ao diretório de genes associados ao termo *language impairment* (*distúrbio da linguagem*, um fenótipo amplo que contempla uma gama variada de comprometimentos linguísticos): <https://hpo.jax.org/app/browse/term/HP:0002463>. Na plataforma, é possível realizar pesquisas mais refinadas a partir de déficits linguísticos específicos (*severe receptive language delay*, por exemplo) e há, para cada gene/distúrbio, *links* externos que direcionam o usuário para artigos cadastrados no *PubMed*, para o OMIM (*Online Mendelian Inheritance in Man*), um catálogo *on-line* que conta com um robusto banco de dados referentes ao genoma humano, apresentando informações detalhadas sobre genes e desordens genéticas (<https://www.omim.org/>) e para o *Orpha.net*, um portal que compila informações sobre doenças genéticas raras, medicamentos e tratamentos (<https://www.orpha.net/consor/cgi-bin/index.php>).

⁷⁸A tabela 4.4, mais uma vez, segue de perto a estabelecida em Burraco (2012).

TABELA 4.4. Distúrbios genéticos que afetam exclusivamente a linguagem (extraído e adaptado de BURRACO, 2012: 250-257)

Gene	Localização cromossômica	Efeitos linguísticos que emergem quando o gene sofre mutação	Denominação clínica do transtorno	Referências bibliográficas relacionadas aos aspectos linguísticos afetados
? (<i>locus SLI1</i>)	16q24	Distúrbio específico de linguagem. O <i>locus</i> aparece vinculado preferencialmente a um comprometimento da memória de trabalho fonológica.	DEL	SLI Consortium (2002); SLI Consortium (2004)
? (<i>locus SLI2</i>)	19q13	Distúrbio específico de linguagem. O <i>locus</i> aparece vinculado a um comprometimento da capacidade expressiva e também a uma dificuldade de repetição de pseudopalavras, indicando um déficit da memória fonológica.	DEL	SLI Consortium (2002); SLI Consortium (2004)
? (<i>locus SLI3</i>)	13q21	Distúrbio específico de linguagem. O <i>locus</i> aparece fortemente vinculado ao endofenótipo do distúrbio caracterizado por um comprometimento da capacidade de leitura.	DEL	Bartlett <i>et al.</i> (2002)
? (<i>locus AUTS1</i>)	7q31	Distúrbio específico de linguagem	DEL	Newbury <i>et al.</i> (2002) ; O'Brien (2003)
?	2p22	Distúrbio específico de linguagem	DEL	Bartlett <i>et al.</i> (2002)
?	17q23	Distúrbio específico de linguagem. O <i>locus</i> aparece debilmente vinculado ao endofenótipo do distúrbio caracterizado por um comprometimento da capacidade de leitura.	DEL	Bartlett <i>et al.</i> (2002)
ATP13A4	3q29	Distúrbio específico de linguagem; atraso significativo na emergência e no desenvolvimento da linguagem.	DEL	Kwasnicka-Crawford <i>et al.</i> (2005)

<i>CACNA1A</i>	19p13	Diferentes graus de afasia (acompanhada de hemiparesia)	Migraña (enxaqueca) hemiplégica familiar	Carrera <i>et al.</i> (1999)
<i>CACNA1C</i>	12p13.3	Desenvolvimento anormal da linguagem (considera-se um transtorno do espectro autista).	Síndrome de Timothy	Splawski <i>et al.</i> (2004)
<i>DYX1C1</i> (<i>locus DDX1</i>)	15q21	Diminuição da capacidade de decodificação fonológica; diminuição da capacidade de decodificação ortográfica; problemas de produção ortográfica; erros na leitura de palavras.	Dislexia	Morris <i>et al.</i> (2000); Taipale <i>et al.</i> (2003)
<i>DCDC2</i> (<i>locus DDX2</i>)	6p22.1	Idem (o <i>locus</i> estaria vinculado a uma variante mais grave de dislexia, atingindo múltiplos componentes fonológicos e ortográficos).	Dislexia	Fisher <i>et al.</i> (1999); Gayan <i>et al.</i> (1999); Grigorenko <i>et al.</i> (2000); Meng <i>et al.</i> (2005); Fisher & Francks (2006); Schumacher <i>et al.</i> (2006)
<i>KIAA0319</i> (<i>locus DDX2</i>)	6p22.2	Idem para as características disléxicas.	Dislexia	Fisher <i>et al.</i> (1999); Gayan <i>et al.</i> (1999); Grigorenko <i>et al.</i> (2000); Marlow <i>et al.</i> (2003); Francks <i>et al.</i> (2004); Cope <i>et al.</i> (2005); Fisher & Francks (2006)
? (<i>locus DDX3</i>)	2p16-p15	Idem	Dislexia	Fagerheim <i>et al.</i> (1999); Kaminen <i>et al.</i> (2003)
? (<i>locus DDX4</i>)	6q11.2-q12	Idem (o <i>locus</i> parece co-segregar com o endofenótipo associado à capacidade de codificação fonológica e à habilidade ortográfica).	Dislexia	Petryshen <i>et al.</i> (2001)

<i>ROBO1 (locus DYY5)</i>	3p12-q13	Idem (o segmento 3p12-q13 também está associado ao SSD [<i>speech-sound disorder</i>], um distúrbio complexo caracterizado por uma dificuldade em usar os sons da fala apropriados à idade devido a déficits articulatorios e/ou fonológicos).	Dislexia; SSD	Stein <i>et al.</i> (2004); Hamula-Jouppi <i>et al.</i> (2005)
? (<i>locus DYY6</i>)	18p11.2	Idem para as características disléxicas.	Dislexia	Fisher & DeFries (2002); Fisher <i>et al.</i> (2002)
<i>DRD4?</i> <i>HRAS?</i> <i>SCT?</i> <i>STIM1?</i> <i>MTR1?</i> (<i>locus DYY7</i>)	11p15.5	Idem (determinadas variantes polimórficas do gene <i>DRD4</i> têm sido correlacionadas com o transtorno de déficit de atenção e hiperatividade [TDAH]; determinadas mutações do gene <i>HRAS</i> têm sido associadas também ao autismo).	Dislexia; TDAH; autismo	Comings <i>et al.</i> (1996); McCracken <i>et al.</i> (2000); Roman <i>et al.</i> (2001); Hsiung <i>et al.</i> (2004)
? (<i>locus DYY8</i>)	1p34-p36	Idem para as características disléxicas.	Dislexia	Rabin <i>et al.</i> (1993); Froster <i>et al.</i> (1993)
? (<i>locus DYY9</i>)	Xq27.3	Idem para as características disléxicas.	Dislexia	De Kovel <i>et al.</i> (2004)
<i>HD (ou HTT)</i>	4p16.3	Graus variados de comprometimento de competências linguísticas de natureza lexical e semântica (nominalização, definição de conceitos, interpretação de significados ambíguos, figurados ou que dependem de inferências) e déficits relacionados à geração de palavras e de orações (fluidez categorial, construção oracional); as maiores dificuldades estão relacionadas à adequada aplicação das regras combinatorias que subjazem à análise morfológica e sintática.	Doença de Huntington	Chenery <i>et al.</i> (2002); Teichman <i>et al.</i> (2005)

<i>GRIK2</i>	6q21	Idem (a presença de determinados polimorfismos no gene influi no aparecimento precoce do transtorno)	Doença de Huntington (fator de risco)	Rubinsztein <i>et al.</i> (1997)
<i>PRNP</i>	20p13	Sintomas semelhantes aos da Doença de Huntington (costuma se manifestar como uma degeneração cognitiva [que em último caso acaba conduzindo à demência], acompanhada de dificuldades de coordenação).	Transtorno neurodegenerativo semelhante à Doença de Huntington (tipo 1); Doença de Creutzfeldt-Jakob (variante hereditária)	Brown <i>et al.</i> (1992); Andrew <i>et al.</i> (1994); Xiang <i>et al.</i> (1998)
<i>JPH3</i>	16q24.3	Sintomas semelhantes aos da Doença de Huntington.	Transtorno neurodegenerativo semelhante à Doença de Huntington (tipo 2)	Margolis <i>et al.</i> (2001); Walker <i>et al.</i> (2002)
? (<i>locus HDL3</i>)	4p15.3	Sintomas semelhantes aos da Doença de Huntington (incluindo uma deterioração particularmente precoce das capacidades cognitivas, perturbação da fala e distonia).	Transtorno neurodegenerativo semelhante à Doença de Huntington (tipo 3)	Al-Tahan <i>et al.</i> (1999)
<i>SRPX2</i>	Xq21-q22	Dificuldade na organização e na coordenação dos movimentos articulatorios orofaciais (na ausência de disartria) e um ligeiro déficit na capacidade perceptiva. Os indivíduos portadores de alguma mutação do gene podem apresentar um padrão anormal de emergência da linguagem e um transtorno da compreensão verbal, que podem vir acompanhados de atraso mental; em outros casos se observa um quadro de disartria ou uma ausência completa da fala, assim como epilepsia e atraso mental (que pode ir de leve a severo).	Epilepsia rolândica com dispraxia verbal (ERDV); polimicrografia perisylviana bilateral (PPB)	Scheffer <i>et al.</i> (1995); Roll <i>et al.</i> (2006).

?	15q14	Erros na geração dos sons da fala, relacionados a problemas na articulação, no processamento fonológico e/ou no processamento linguístico.	Transtorno dos sons da fala (SSD: <i>speech-sound disorder</i>)	Stein <i>et al.</i> (2006)
	<i>ProSAP2 (SHANK3) / FLJ10659</i> 22q13.3/ 12q24.1	Atraso severo na emergência da linguagem, que pode estar completamente ausente em alguns casos.	Síndrome de Phelan-McDermid	Prasad <i>et al.</i> (2000); Bonaglia <i>et al.</i> (2001); Phelan <i>et al.</i> (2001); Monteiro & Ferg (2017)

RESUMO DO CAPÍTULO E CONSIDERAÇÕES FINAIS

Ao longo deste capítulo, discutimos mais detalhadamente alguns fatores genéticos (possivelmente) associados à linguagem. Inicialmente, vimos como determinados modelos genéticos teóricos (como o de Hoffman, 2003) podem oferecer ferramentas eficientes de investigação biolinguística. A ideia central por trás de tais modelos é a de que toda atividade linguística corresponde a uma atividade neuronal que, por sua vez, é fruto de uma atividade gênica subjacente responsável pelo controle dos processos metabólicos que dão origem à linguagem. Assim, cada fenômeno linguístico emerge no cérebro como o resultado de genes que se expressam diferencialmente nos núcleos neuronais. Alguns destes genes controlariam a construção e o desenvolvimento da estrutura geral dos núcleos linguísticos – sendo responsáveis pela construção de neurônios especializados em processar informação linguística (especialização *a priori*, anterior ao *input* linguístico) ou, por outra via, sendo responsáveis pela especialização linguística de neurônios dotados de plasticidade funcional (especialização *a posteriori*, a partir do *input* linguístico) – enquanto outros atuariam como reguladores do processamento em tempo real de tarefas linguísticas, expressando-se de forma diferencial de acordo com a demanda informacional exigida, i.e., de acordo com o tipo de *puzzle* linguístico a ser computado e resolvido (PINKER, 1994). Em tese, portanto, o mapeamento de possíveis fatores genéticos ligados à linguagem poderia ser realizado a partir da identificação e do cálculo de taxas diferenciais de produtos gênicos expressos durante a realização da linguagem no cérebro.

Apesar da tecnologia de que dispomos hoje – que é capaz de medir os níveis de atividade metabólica do cérebro durante a realização de tarefas linguísticas, indicando em quais áreas a atividade gênica é maior –, a forma mais utilizada para se encontrar os fatores genéticos responsáveis por uma determinada competência linguística ainda é a que consiste em identificá-la e mapeá-la a partir de suas variantes fenotípicas desviantes. Por isso, os estudos genéticos sobre linguagem partem tradicionalmente da caracterização e da análise de patologias que afetam de alguma forma a linguagem. O objetivo central desse tipo de abordagem é encontrar a contraparte genética de uma determinada disfunção linguística através do mapeamento de erros em sequências específicas de DNA, identificando, assim, os fatores (genes) responsáveis pela competência linguística afetada pelo distúrbio.

É relativamente comum que problemas de linguagem não apareçam de forma isolada, mas façam parte de síndromes ou doenças genéticas caracterizadas por um quadro amplo de

sintomas. Alguns tipos de erros inatos do metabolismo, por exemplo, acabam fazendo com que os indivíduos afetados manifestem uma série de problemas de saúde, incluindo disfunções linguísticas mais ou menos graves, que muitas vezes são derivadas de problemas subjacentes (como disfunções na comunicação entre os neurônios) ou integram quadros patológicos mais amplos (como atraso global do desenvolvimento). Nestes casos, mutações em determinados genes acabam gerando uma série de efeitos deletérios em inúmeros processos fisiológicos, atingindo, inclusive, competências neurocognitivas como a linguagem.

Por fim, discutimos os aspectos ontogenéticos de alguns subtipos de DEL – uma categoria de distúrbios que congrega uma série de transtornos de natureza exclusivamente linguística. Não raramente, casos de DEL afetam membros de uma mesma família (agregação familiar significativa) e apresentam taxas de concordância elevadas entre gêmeos idênticos, o que aponta para a existência de fatores hereditários/genéticos responsáveis pelo condicionamento de determinados padrões neurológicos que podem servir de substrato para a desencadeamento do distúrbio, cuja manifestação concreta depende de uma junção de fatores genéticos, epigenéticos e ambientais (causa multifatorial).

Ainda neste capítulo, fizemos uma pequena reformulação no enunciado do problema que tem nos orientado até aqui – encontrar as possíveis relações entre genes e linguagem –, a fim de adaptá-lo às informações disponíveis no campo dos distúrbios linguísticos. O novo enunciado – *qual é (ou quais são) o(s) efeito(s) linguístico(s) provocado(s) pela mutação de um determinado gene?* – mantém a essência do problema e nos oferece um ponto de partida concreto (gene mutado $X \rightarrow$ efeito linguístico Y) a partir do qual podemos proceder a análises teóricas mais avançadas que busquem relacionar, por exemplo, genes e competências (gene $X \rightarrow$ competência linguística A).

CAPÍTULO 5 – *FOXP2*

Como o FOXP2 foi descoberto? Qual é o seu papel biológico? De que modo ele se expressa no cérebro e de que forma ele pode estar associado à linguagem? Pode um gene que se expressa no coração de humanos, no cérebro de canários e em neandertais ter realmente algum papel na determinação de propriedades linguísticas? Quais foram os efeitos causados por uma mutação do FOXP2 em uma família em que metade dos seus membros apresentava um distúrbio grave de fala e linguagem?

Antes de iniciarmos a nossa discussão sobre o *FOXP2*, vamos revisar algumas ideias centrais.

1. Ao todo, o genoma humano contém aproximadamente três bilhões de pares de nucleotídeos (LANDER *et al.*, 2001; VENTER *et al.*, 2001), distribuídos ao longo de 46 cromossomos. No entanto, a porção funcional de DNA (i.e., aquela que pode ser expressa em produtos gênicos) é muito pequena e representa menos de 3% do genoma. Ou seja, dos três bilhões de pares de nucleotídeos que formam o programa genético humano, apenas uma pequena fração tem função biológica definida e faz parte dos aproximadamente 30.000 genes que compõem a nossa espécie. A maior parte das sequências nucleotídicas não codifica proteínas e compõe o chamado DNA não codificante (ou DNA intergênico), cujas funções estão relacionadas a mecanismos regulatórios (síntese de RNA não codificante, regulação transcricional etc.). Assim, se existe uma matriz genética responsável por tornar os humanos aptos a adquirir e desenvolver uma língua, é nesse conjunto reduzido de DNA codificante que ela se encontra.
2. Um gene é um segmento de DNA responsável pela síntese de uma proteína (mais precisamente, de uma cadeia polipeptídica). No núcleo das células dos seres eucarióticos (aqueles cujas células possuem um núcleo delimitado, separado do citoplasma por uma membrana), o DNA que compõe o gene é transcrito em uma molécula de RNA mensageiro (RNAm). Após ser editado (em um processo

denominado *splicing*), o RNAm atravessa a membrana nuclear e, no citoplasma, conecta-se aos ribossomos, que traduzem a informação codificada, dando origem à cadeia polipeptídica/proteína. Após a tradução, as informações codificadas nos genes se transformam em materiais biológicos. Neste sentido, os genes funcionam como receitas moleculares, pois contêm as instruções para a construção e o funcionamento da vida. Como veremos mais adiante, o *FOXP2* é responsável pela produção de uma proteína que funciona como fator de transcrição (uma classe especial de proteínas que participa da ativação/inativação de outros genes).

O mecanismo central da biologia molecular



3. No núcleo celular, o DNA está subdividido em filamentos chamados cromossomos. Cada cromossomo possui um conjunto específico de genes (além de conter também proteínas e DNA não codificante). A espécie humana possui 46 cromossomos: 44 autossomos e 2 cromossomos sexuais (também chamados de alossomos). Durante a gametogênese, as células germinativas humanas se dividem, dando origem a gametas com um lote de 23 cromossomos ($n = 23$). Na fecundação, os gametas masculino e feminino se fundem, formando o zigoto. Desta forma, o organismo gerado recebe dois lotes cromossômicos ($2n = 46$) e possui duas cópias para cada um dos cromossomos. Consequentemente, cada célula somática de seu corpo possuirá duas versões (dois alelos) de um mesmo gene (uma localizada no cromossomo de origem materna e a outra localizada na região homóloga do cromossomo de origem paterna), que poderão ou não ser bioquimicamente iguais. Esta distribuição de genes em pares homólogos determina os dois tipos básicos de distúrbios genéticos que podem derivar da relação entre os alelos: distúrbios dominantes e distúrbios recessivos. Nos distúrbios dominantes, apenas uma cópia mutada do gene é suficiente para determinar o transtorno (ou seja, o alelo mutante dominante $[A]$ determina o fenótipo mesmo quando se encontra em dose simples, em indivíduos heterozigóticos $[Aa]$). Este é o padrão de dominância observado no distúrbio causado por mutações no *FOXP2*. Já nos distúrbios recessivos, são necessárias duas cópias mutadas do gene para

desencadear o transtorno (ou seja, o alelo mutante recessivo [a] só determina o fenótipo quando está em dose dupla, em indivíduos homozigóticos recessivos [aa]).

4. Após a fecundação, o zigoto passa a se dividir e a se diferenciar, apresentando um nível crescente de especialização de suas células, que passam de um estado inicial totalmente indiferenciado e totipotente – onde são capazes de dar origem a qualquer tipo celular – para um estado diferenciado, quando assumem forma e função específicas, organizando-se em tecidos e em estruturas cada vez mais complexas. Esse processo de especialização celular é orquestrado por um conjunto de sinais (hormônios e outras substâncias da matriz extracelular, por exemplo) que, ao se conectarem à membrana celular de uma determinada célula, provocam uma reação em cadeia no meio intracelular, culminando com a ativação ou a repressão de certos genes. Essa configuração específica de genes ativados gera uma combinação específica de proteínas sintetizadas (proteoma) que será responsável pela especialização morfológica e funcional da célula, fazendo com que ela deixe de ser um sistema indiferenciado que executa inúmeras funções com baixa eficiência e passe a ser um sistema otimizado capaz de executar funções específicas com alta eficiência.
5. O processo de diferenciação celular e de desenvolvimento inicial dos organismos está intimamente ligado ao funcionamento de uma classe especial de genes regulatórios (genes homeóticos) que têm a função de controlar os padrões básicos de distribuição celular e de desenvolvimento corporal do organismo, definindo os locais precisos em que cada estrutura deve se diferenciar e se desenvolver. Os genes desta classe – que, por compartilharem uma sequência específica de DNA chamada *homeobox*, foram agrupados em uma mesma família (família *Hox*) – são responsáveis pela produção de fatores de transcrição específicos (proteínas especializadas) que se ligam a determinados genes, ativando-os em fases específicas do desenvolvimento embrionário. Estes genes “receptores”, por sua vez, também podem sintetizar outros fatores de transcrição, que serão responsáveis pela regulação de outros genes, numa espécie de efeito em cascata altamente organizado que, ao fim e ao cabo, acaba dando origem aos diferentes tipos celulares e às diferentes estruturas/regiões corporais que compõem o organismo em desenvolvimento. Outras classes de genes – como a que inclui o *FOXP2* (família *FOX*) – também são responsáveis pela síntese de fatores de transcrição que atuam diretamente no controle da atividade transcriptiva de outros

genes, modulando taxas diferenciais de produção proteica e permitindo que células que compartilham o mesmo material genético possam se diferenciar morfológica e funcionalmente umas das outras devido aos conjuntos distintos de proteínas que estão sendo sintetizadas em seu interior.

6. Cada ser humano que existe ou que já existiu no planeta possui um genoma único (com exceção dos gêmeos idênticos, que, mesmo assim, podem possuir pequenas diferenças em seu DNA em consequência de mutações que ocorrem após a divisão do embrião inicial em dois ou mais embriões derivados). O fato de sermos únicos, no entanto, não anula o fato de sermos extremamente semelhantes. Dois genomas humanos – de quaisquer etnias, espaços geográficos ou tempos históricos – terão sempre uma taxa de concordância de 99,9% (aproximadamente) e serão praticamente idênticos do ponto de vista genético. A explicação para esse aparente paradoxo (genomas tão semelhantes dando origem a seres tão singulares) está na extensão do nosso programa genético, uma vez que o espaço de variação entre dois indivíduos (0,1%) ainda deixa margem para milhões de diferenças acumuladas entre os 3 bilhões de pares de nucleotídeos que compõem o nosso genoma. É nesse espaço de variação (relativamente reduzido, por um lado, e relativamente amplo, por outro) que estão alocadas as diferenças genotípicas que darão origem às diferenças fenotípicas entre dois indivíduos quaisquer. Assim, se existe uma matriz em nosso DNA responsável pela linguagem, espera-se que os indivíduos afetados por problemas linguísticos geneticamente condicionados apresentem diferenças em determinadas partes de seu genoma quando comparados a indivíduos não afetados. De certa forma, a história do *FOXP2* sintetiza essa tentativa de buscar nas profundezas do DNA erros que possam não apenas explicar determinados transtornos linguísticos, mas sobretudo fornecer informações sobre a natureza biológica da linguagem.

5.1. BREVE HISTÓRIA DE UMA DESCOBERTA

No início dos anos 90, uma família britânica – que entraria para a história da Linguística como família KE – começou a chamar a atenção da comunidade científica internacional devido ao número elevado de membros que apresentavam um distúrbio caracterizado, sobretudo, por uma série de comprometimentos linguísticos de graus variados. Como as principais dificuldades dos membros afetados estavam relacionadas à coordenação dos movimentos da fala, o distúrbio foi inicialmente classificado como uma dispraxia verbal do desenvolvimento (HURST *et al.*, 1990). Como já mencionamos anteriormente, nas apraxias e dispraxias verbais, os indivíduos não conseguem programar e executar os movimentos articulatorios necessários à produção adequada dos fonemas durante a fala. De fato, todos os membros afetados da família apresentavam limitações orofaciais severas e sua fala era considerada muitas vezes incompreensível por ouvintes da própria família. Em estudos posteriores (GOPNIK, 1990; GOPNIK & CRAGO, 1991), no entanto, foram identificados déficits especificamente linguísticos, como dificuldades na pluralização de pseudopalavras (mas não na pluralização de itens lexicais conhecidos pelos falantes, que lançavam mão de estratégias de memorização para expressar o plural, indicando uma preservação do sistema mnemônico de acesso lexical) e no processamento de traços gramaticais de tempo e pessoa, o que sugeriu a hipótese de um déficit no sistema morfossintático da gramática internalizada dos membros afetados. Aparentemente, os indivíduos afetados memorizavam as formas flexionadas e conjugadas, mas não conseguiam aplicar, de modo espontâneo (“instintivo”), determinadas regras de formação morfológica. Uma palavra como *books*, por exemplo, não era processada pelos membros afetados como o resultado de uma regra geral de pluralização aplicada à palavra *book*, mas era memorizada como um item indecomponível que continha em si mesmo a noção de plural (ou seja, ao invés de executarem uma regra de sufixação do tipo $[_N \text{ book}] + [_{pl.} -s] \rightarrow [_N \text{ books}]$, os membros afetados, aparentemente, possuíam duas entradas lexicais em seu dicionário mental, uma para o item *book* e outra para o item *books*). Discutiremos mais detalhadamente esses e outros déficits linguísticos da família KE na seção 5.4.

Outras análises, porém, constataram que o distúrbio não afetava apenas aspectos específicos da gramática dos falantes (Fig. 5.1a), mas era antes caracterizado por uma degeneração mais ampla do sistema expressivo, incluindo defectividades articulatorias severas que confirmavam a existência de uma dispraxia verbal (VARGHA-KHADEM & PASSINGHAM, 1990; VARGHA-KHADEM *et al.*, 1995). Adicionalmente, foram

apresentadas evidências de que os indivíduos afetados apresentavam problemas subjacentes na coordenação da musculatura orofacial responsável não apenas pela execução mecânica da fala, mas também por movimentos não linguísticos, caracterizando uma dispraxia orofacial (Fig. 5.1b). Ambas as dispraxias – orofacial e verbal – estavam presentes em todos os membros afetados da família e apareciam precocemente na primeira infância, persistindo na idade adulta. Além disso, os indivíduos afetados apresentavam valores em testes de cognição um pouco abaixo do normal, com déficits mais pronunciados no domínio verbal (FISHER *et al.*, 1998).

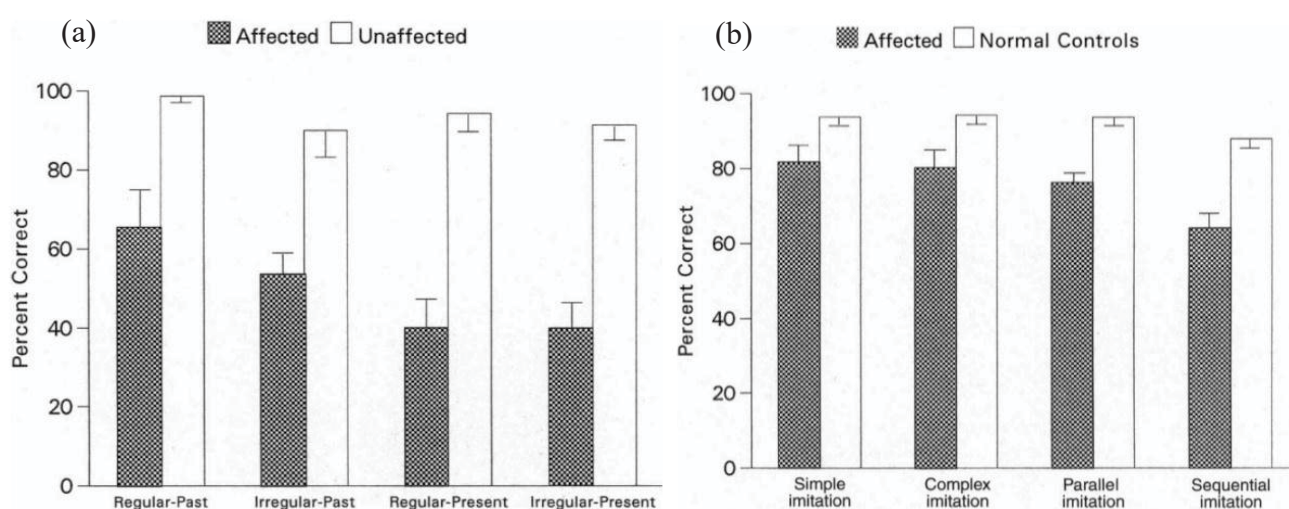


FIGURA 5.1. (a) Aplicação de regras flexionais e produção de tempos verbais. Os participantes do experimento – todos da família KE – eram solicitados a completar adequadamente a lacuna verbal de uma determinada sentença a partir da forma verbal utilizada na sentença antecedente, fazendo, para isso, as modificações flexionais necessárias. Assim, por exemplo, quando o sujeito tinha que gerar uma forma flexionada no passado, propunham-se sentenças do tipo “*Every day I wash my clothes; yesterday I _____ my clothes*” [washed]. De forma semelhante, quando o sujeito precisava gerar uma forma flexionada no presente, propunham-se exercícios do tipo “*Yesterday I washed my clothes; every day I _____ my clothes*” [wash]. Os resultados mostram desníveis importantes na capacidade de produção gramatical dos tempos verbais entre indivíduos afetados e não afetados da família KE. **(b)** Imitação de movimentos orais e faciais. Os sujeitos participantes do experimento foram solicitados a imitar vários tipos de movimentos: movimentos simples (como abrir a boca), movimentos complexos (como avançar a mandíbula e o lábio), séries de três movimentos realizados em paralelo (como abrir a boca enquanto estica a língua e realiza uma vocalização) e séries de três movimentos sequenciais (como fechar os lábios, para depois abrir a boca e colocar a língua para fora). Os resultados mostram que, quanto maior é o grau de dificuldade do movimento, maiores são os desníveis entre o desempenho dos indivíduos afetados da família KE e o desempenho dos indivíduos não afetados do grupo controle. Dada a complexidade dos movimentos relativos à execução da fala, não é de se estranhar que os KE afetados apresentem dificuldades em coordenar a musculatura necessária à produção sequencial de vogais, consoantes e sílabas. Note-se, entretanto,

que a assimetria entre afetados e não afetados observada no teste linguístico é mais expressiva do que a assimetria identificada no teste motor. Gráficos extraídos de VARGHA-KHADEM *et al.* (1995).

Posteriormente, Watkins *et al.* (2002) mostraram que alguns comprometimentos apresentados pelos KE não poderiam ser explicados apenas como uma consequência do déficit no controle motor orofacial. Foram observadas, por exemplo, dificuldades com relação à repetição de palavras (reais ou não) e à manipulação de fonemas também no nível escrito, indicando um problema processual de base. Além disso, estes autores identificaram que os indivíduos afetados, além de manifestarem problemas em termos de produção linguística, também apresentavam déficits em termos de compreensão, apresentando resultados ruins em tarefas de decisão lexical (nas quais os pacientes precisavam fazer a distinção entre palavras reais e inventadas). Dificuldades com relação à compreensão também foram diagnosticadas por Fisher *et al.* (2003) no nível morfo sintático, demonstrando que os KE afetados apresentavam dificuldades tanto para a geração de estruturas (através da flexão ou derivação) quanto para sua compreensão, com um comprometimento especial relacionado ao processamento de sentenças complexas (FISHER *et al.*, 2003).

Além da riqueza do fenótipo e do número elevado de membros familiares afetados pelo transtorno, o que chamou realmente a atenção dos pesquisadores foi o padrão de herança apresentado pelo distúrbio. Ainda em 1987, alguns membros da família KE já haviam sido encaminhados para um estudo de aconselhamento genético devido à alta incidência de problemas relacionados à fala. Ao se analisar a genealogia da família (Fig. 5.2), comparando-se indivíduos afetados com indivíduos não afetados de três gerações consecutivas, descobriu-se um padrão hereditário autossômico (ou seja, comum a ambos os sexos) dominante com relação à distribuição do transtorno de fala e linguagem. Aparentemente, apenas um alelo mutante dominante estava condicionando o distúrbio, correspondendo a um caso de herança monogênica (ou seja, relacionada a um único gene) – algo relativamente raro, já que fenótipos associados à cognição tendem a apresentar padrões complexos de herança multifatorial, sendo determinados por um conjunto de genes que interagem entre si sob a influência de fatores ambientais.

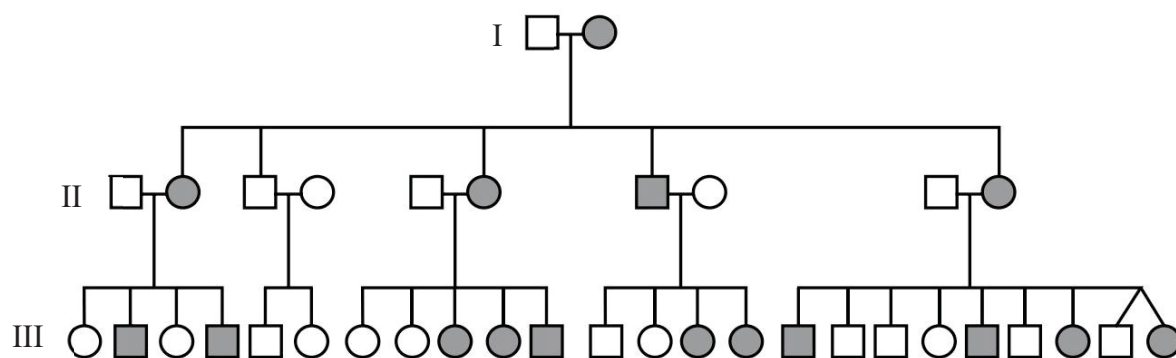


FIGURA 5.2. Árvore genealógica da família KE. Homens são representados por quadrados; mulheres são representadas por círculos. Os membros familiares afetados pelo distúrbio de fala e linguagem estão representados por símbolos escuros e são todos heterozigóticos (Ff) para o gene de interesse. Os demais são homozigóticos recessivos (ff) e não são afetados. (I), (II), e (III) indicam as gerações, da mais velha (I) para a mais nova (III). Heredograma extraído e adaptado de LEONARD (2014).

A configuração da árvore genealógica da Figura 5.2 tornava evidente a etiologia genética do distúrbio apresentado pelos KE. A “mutação original” foi identificada na avó da geração (I). Os indivíduos afetados eram heterozigóticos (Ff) e transmitiam aos descendentes uma cópia não funcional (F) ou uma cópia funcional (f) do gene. O distúrbio, inicialmente, não parecia ser influenciado pelo meio ambiente, uma vez que os indivíduos conviviam sob as mesmas condições gerais de existência: pertenciam à mesma família, recebiam a mesma educação e compartilhavam do mesmo ambiente linguístico, estando, teoricamente, expostos ao mesmo conjunto de estímulos. Apesar de cada indivíduo vivenciar um conjunto único de experiências capazes de moldar determinados aspectos de sua cognição, seria especialmente difícil explicar como (e quais) fatores ambientais poderiam estar desencadeando um distúrbio grave de fala e linguagem que afetava quase metade dos indivíduos de uma mesma família, preservando a outra metade. Todas as evidências, portanto, apontavam para a existência de um gene mutado responsável pelo distúrbio da família KE.

Mesmo que o padrão monogênico do distúrbio apresentado pelos KE seja um objeto de estudo menos intrincado do que o padrão complexo de fenótipos causados pela interação de múltiplos genes, encontrar um gene específico entre bilhões de bases nitrogenadas não é uma tarefa simples. A fim de delimitar o campo de busca, os pesquisadores usam uma série de técnicas de rastreamento, incluindo o uso de marcadores genéticos, que são sequências de nucleotídeos de localização conhecida nos cromossomos e que servem como pontos estratégicos de referência. A ideia central desse tipo de análise é comparar o padrão de

segregação de um marcador genético – i.e., a forma como ele se distribui ao longo dos membros de uma mesma família – com o padrão de segregação de um determinado fenótipo, a fim de identificar correlações. Se ambos os padrões coincidirem, temos um indício de que aquela sequência de DNA usada como marcador pode estar próxima do gene de interesse. Se o marcador genético e o traço fenotípico apresentarem padrões distintos de segregação, temos um forte indício de que o gene responsável pela característica analisada não está nas proximidades do trecho de DNA que foi usado como marcador. Assim, se um marcador genético X for herdado por todos os membros afetados de uma família e não estiver presente nos indivíduos não afetados, é provável que o gene responsável pelo condicionamento do fenótipo analisado esteja localizado no interior ou nas proximidades desse marcador X.

Essa análise comparativa entre padrões de segregação se baseia no princípio de que quanto mais próximas estão duas sequências de DNA em um cromossomo, maiores serão as chances de elas serem herdadas conjuntamente. Os genes, em um cromossomo, ocupam posições específicas e se encontram ligados uns aos outros (*linkage*). Durante uma determinada fase da formação dos gametas (mais especificamente, na prófase I da meiose), os pares de cromossomos homólogos se emparelham (sinapse) e trocam trechos de DNA entre si (*crossing over* ou permutação), adquirindo novas configurações de genes (Fig. 5.3). Nesse processo de *recombinação gênica*, quanto maior for a proximidade entre dois genes (ou duas sequências quaisquer de DNA), menores serão as chances de que o “ponto de corte” (quiasma) ocorra em uma região entre os dois, que tendem a permanecer juntos no mesmo cromossomo, sendo herdados conjuntamente pelo futuro zigoto. Desta forma, se o padrão de herança de um determinado fenótipo coincide com o padrão de herança da sequência de DNA que é usada como marcador genético, há uma grande possibilidade de que o gene responsável pelo traço fenotípico investigado esteja nas proximidades do marcador, já que trechos circunvizinhos de DNA tendem a permanecer ligados.

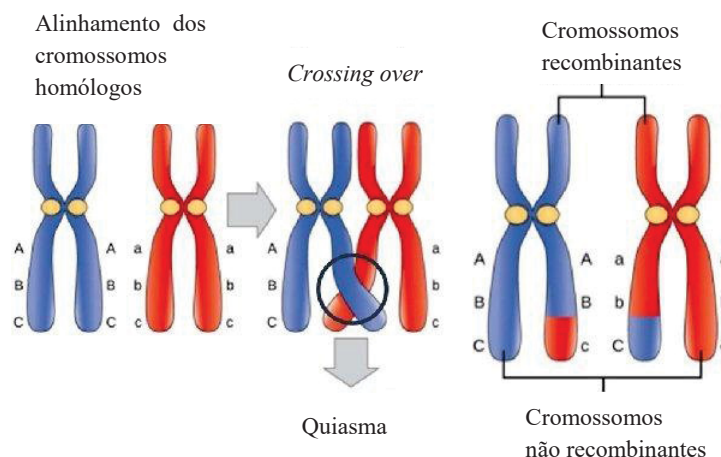


FIGURA 5.3. *Crossing over* (permutação) e recombinação gênica. Na prófase I da meiose, o emparelhamento dos cromossomos (que já estão duplicados e por isso formam um “X”) pode resultar em quebra do DNA em regiões homólogas, seguida de troca de trechos cromossômicos. Cromossomos recombinantes são os que apresentam as novas combinações gênicas resultantes da permutação (ABc e abC); cromossomos não recombinantes (ou parentais) são os que mantêm as combinações gênicas originais (ABC e abc). Quanto mais próximos estão dois trechos de DNA menor é o risco de separação e de recombinação durante a gametogênese. No exemplo da figura, o *locus* dos alelos *A* e *a* se mantém ligado ao *locus* dos alelos *B* e *b* (mais próximo), mas sofre permutação com relação ao *locus* dos alelos *C* e *c* (mais distante). Imagem extraída e adaptada de CLARK *et al.* (2018). Disponível em: <https://openstax.org/books/biology-2e/pages/11-1-the-process-of-meiosis>

No caso dos distúrbios genéticos, a presença de um marcador molecular específico em todos os indivíduos afetados não licencia o estabelecimento de uma relação imediata de causa e efeito entre aquela sequência de DNA utilizada como marcador e o fenótipo investigado – isto porque a mutação responsável pelo distúrbio pode estar localizada em sequências nucleotídicas adjacentes ao marcador e não necessariamente em seu interior. Assim, espera-se que a mutação procurada esteja nas imediações do marcador genético, que funciona como um ponto de referência molecular dentro do cromossomo. Um passo importante na investigação do fator genético responsável pelo distúrbio dos KE, portanto, consistia em encontrar algum marcador genético cujo padrão de segregação coincidisse com o do distúrbio de linguagem. Em 1998, Fisher *et al.* anunciaram a identificação de um marcador genético com tais características (Fig. 5.4). Localizado na região 7q31 do cromossomo 7, os pesquisadores deram ao *locus* o nome de *SPCHI*, numa alusão ao termo *speech* (fala; discurso). O próximo passo, então, consistia em pesquisar esse *locus* – que media aproximadamente

5,6 cM⁷⁹ e que continha mais de 70 genes conhecidos – a fim de identificar qual dos genes localizados nessa região havia sido danificado, começando a varredura por aqueles que já haviam sido relacionados de alguma forma a funções neurocognitivas.

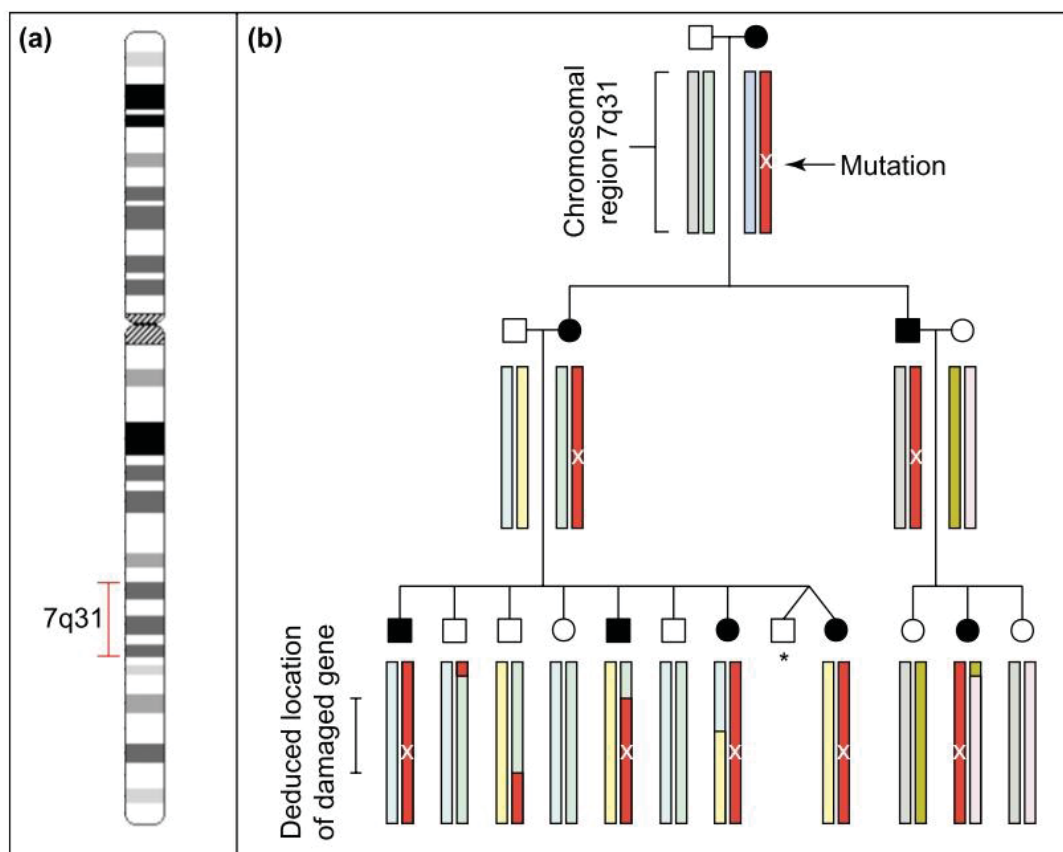


FIGURA 5.4. Localizando o gene danificado na família KE. Para simplificar, apenas um subconjunto da família é mostrado em (b), incluindo dois ramos contendo 8 do total de 15 indivíduos afetados. Quadrados representam homens, círculos representam mulheres; indivíduos sombreados em preto são afetados pelo distúrbio de fala e linguagem, indivíduos não sombreados são normais. A estrela indica um gêmeo não idêntico cujo DNA não estava disponível para estudo. Todo ser humano herda duas cópias de cada cromossomo (uma paterna e uma materna), e seções dessas cópias são embaralhadas de geração em geração por um processo de troca conhecido como recombinação gênica. Assim, a composição genômica que é herdada por um indivíduo é melhor vista como uma colcha de retalhos dos cromossomos maternos e paternos de gerações anteriores, e isso pode ser explorado para localizar genes danificados. Os geneticistas usaram informações de uma série de marcadores genéticos para rastrear o padrão de herança de regiões cromossômicas através da família KE. Eles encontraram resultados particularmente interessantes para uma seção do cromossomo 7, que correspondia a uma parte da banda cromossômica conhecida como 7q31 (a). Em (b) cores diferentes indicam as origens de diferentes cópias

⁷⁹ Centimorgan (cM): unidade de mapeamento genético utilizada para medir distâncias relativas dentro do cromossomo. Um centimorgan corresponde a cerca de 1 milhão de pares de bases (pb) nitrogenadas em humanos. $10^6 \text{ pb} = 1 \text{ Mb}$ (1 mega de pares de bases). $\sim 5,6 \text{ cM} = \sim 5,6 \cdot 10^6 \text{ pb} = \sim 5,6 \text{ Mb}$ (medida relativa ao genoma humano).

de cromossomos dentro da família KE. Descobriu-se que todos os indivíduos afetados na família KE herdaram uma cópia idêntica de uma porção de 7q31. Esta cópia se originou de um cromossomo da avó afetada (marcada em vermelho). Ao determinar a porção precisa da cópia marcada em vermelho de 7q31 que era comum a todos os membros afetados (mas não herdada por nenhum membro não afetado), foi possível deduzir a localização provável do gene danificado. Os pesquisadores foram então capazes de se concentrar em genes desta porção específica do cromossomo 7 em seus esforços subsequentes de triagem da mutação, conseguindo, posteriormente, encontrá-la (a mutação está indicada aqui por um X branco). Extraído e adaptado de MARCUS & FISHER (2003).

A descoberta casual de outro indivíduo (um menino, identificado por CS) acelerou o processo de identificação do gene investigado. Testes linguísticos e neurológicos haviam identificado em CS déficits muito similares aos dos KE afetados. Aos dois anos de idade, CS começou a apresentar um leve distúrbio motor e um atraso no desenvolvimento da fala; aos três anos e meio foi diagnosticado com dispraxia verbal, que se manifestava na forma de um distúrbio na compreensão e na execução do discurso falado, fazendo com que ele atingisse a fase de duas palavras apenas aos quatro anos de idade (LAI *et al.*, 2000). Além disso, sabia-se que CS apresentava uma translocação (t) entre os cromossomos 5 e 7. Nas translocações cromossômicas, os cromossomos se quebram e são “ressoldados” de modo cruzado. Neste caso, uma parte do braço longo (q) do cromossomo 5 de CS havia sido transferida para o braço longo (q) do cromossomo 7 e vice-versa: t(5;7)(q22; q31.2). Ao analisarem o cromossomo 7 de CS, os pesquisadores descobriram que ele havia se rompido em uma região interna a *SPCH1*.

Uma análise mais precisa mostrou que esse ponto de interrupção havia atingido diretamente um gene que pertencia a um grupo de genes reguladores responsáveis pela codificação de proteínas *forkhead* – uma classe de fatores de transcrição que desempenham uma série de funções na diferenciação e proliferação celular, na formação de padrões morfológicos e na transdução de sinal (mecanismo através do qual uma molécula de sinalização extracelular ativa um receptor específico localizado na membrana celular, desencadeando uma cascata bioquímica de eventos intracelulares que gera uma resposta ao estímulo inicial). Muitos destes fatores de transcrição estão implicados na regulação do desenvolvimento do embrião em crescimento e alguns também têm papéis distintos nos tecidos adultos, atuando, por exemplo, no controle do metabolismo (MARCUS & FISHER, 2003). Para entender como os fatores de transcrição atuam é preciso lembrar que um gene é constituído por uma região codificadora – responsável por armazenar a “receita” molecular de uma determinada proteína, contendo sequências de nucleotídeos que serão transcritas em

moléculas de RNA mensageiro e traduzidas em sequências de aminoácidos – e uma região reguladora, cuja função principal é determinar, a cada momento, quantas cópias de RNA devem ser feitas pela maquinaria de transcrição em função das demandas da célula. Fatores de transcrição como as proteínas *forkhead* são elementos que interagem com as regiões reguladoras de outros genes, acionando ou inibindo sua atividade transcriptiva. Essa propriedade faz com que os fatores de transcrição tenham um papel fundamental na modulação do nível de expressão de outras proteínas nas células, controlando sua abundância e disponibilidade. A característica unificadora de todas as proteínas *forkhead* é a **Forkhead BOX** (daí a designação proteínas/genes *FOX*), uma pequena cadeia de 80-100 aminoácidos que forma um domínio de ligação capaz de se conectar à região reguladora de outros genes. É esse domínio funcional que dá às proteínas *FOX* a propriedade de controlar a expressão dos genes-alvo. No caso do menino CS, a quebra do cromossomo 7 havia atingido em cheio o gene *FOX* (Fig. 5.5) da região SPCH1, que recebeu o nome de *FOXP2*.

A análise do *FOXP2* nos indivíduos afetados pertencentes à família KE revelou que todos eles possuíam uma mutação pontual na sequência codificadora do gene – mais especificamente, um polimorfismo de nucleotídeo único (SNP) que havia provocado a substituição de uma única guanina (G) por uma adenina (A) no éxon 14 do *FOXP2* (Fig. 5.5) –, o que resultava na troca de um aminoácido na sequência da proteína FOX sintetizada (mais precisamente, a substituição de uma arginina (Arg), um aminoácido essencial para a ligação da proteína à região reguladora de outros genes, por uma histidina (His) não funcional [LAI *et al.*, 2001]). Essa mutação (classificada como R553H), que resulta na troca de um único aminoácido em uma região nobre da proteína, presumivelmente provoca uma mudança na carga eletrostática do seu domínio de ligação FOX, o que a impede de se ligar corretamente ao DNA de outros genes (de fato, trabalhos posteriores demonstraram que essa substituição impede que proteínas da família FOX consigam se conectar ao DNA *in vitro* [VERNES *et al.*, 2006]). Sem essa capacidade de se unir aos genes-alvo, a proteína codificada pelo *FOXP2* perde a capacidade de atuar como fator de transcrição, tornando-se disfuncional (Fig. 5.6). Além de apresentar o mesmo padrão de segregação identificado nos indivíduos da família KE que eram afetados pelo distúrbio, a troca de uma guanina por uma adenina no éxon 14 do *FOXP2* não estava presente em nenhum dos 364 cromossomos de indivíduos não afetados que foram utilizados como grupo controle. Assim, com o conjunto acumulado de evidências relacionadas à família KE e a CS, estabelecia-se pela primeira vez uma ligação inequívoca entre a mutação de um gene e um distúrbio do desenvolvimento da fala e da linguagem.

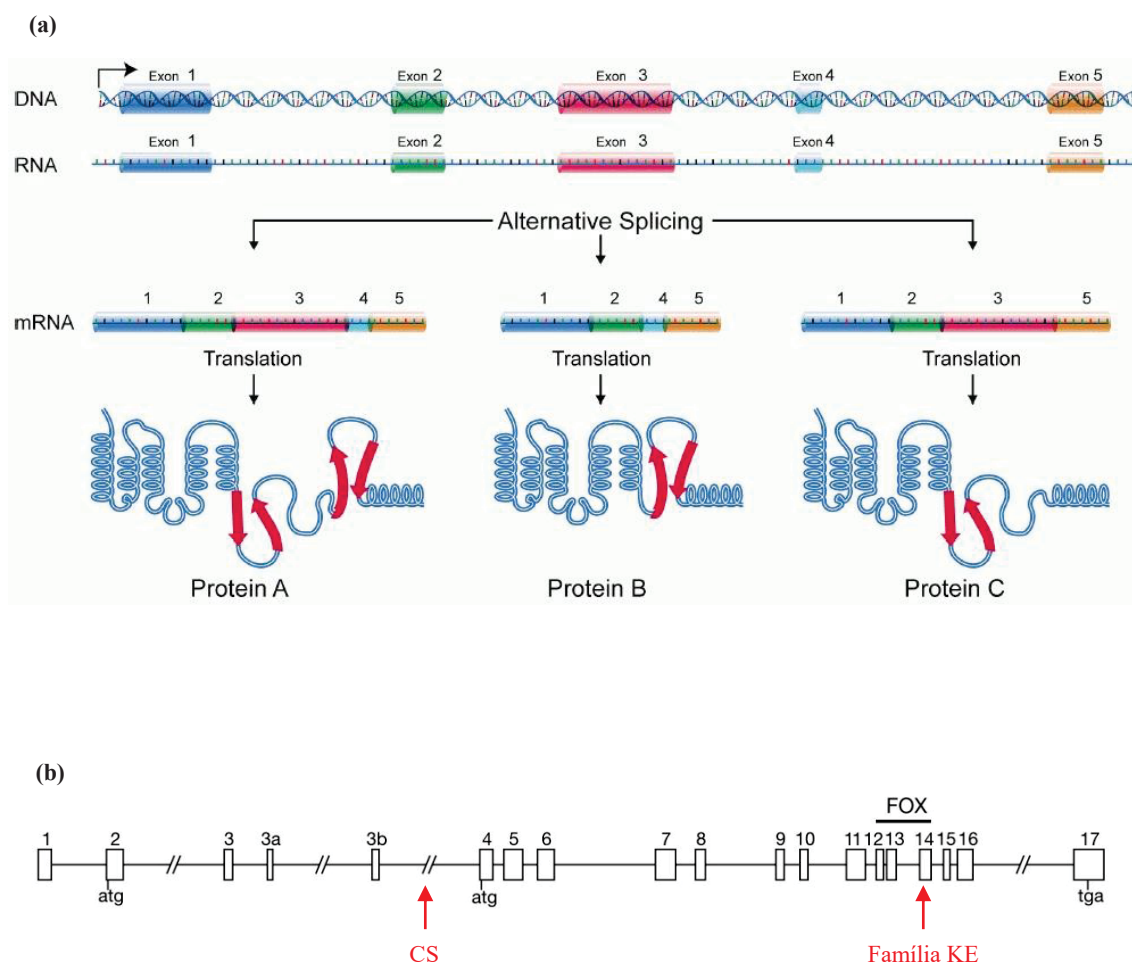


FIGURA 5.5. (a) Transcrição e *splicing* alternativo. No *splicing*, algumas seções não codificadoras do RNA transcrito (íntrons: *intragenic regions*) são removidas, e as seções restantes (éxons: *expressive regions*) são acopladas novamente. Alguns genes podem sofrer *splicing* alternativo, levando à produção de diferentes moléculas maduras de RNAm a partir do mesmo transcrito inicial. Cada versão alternativa do RNAm é traduzida em uma proteína com propriedades estruturais e funcionais particulares. Fonte da imagem (a): https://commons.m.wikimedia.org/wiki/File:DNA_alternative_splicing.gif. Acesso em 22/12/2021. **(b)** Representação simplificada da estrutura do gene *FOXP2* humano. Os retângulos numerados representam os éxons, com as posições dos códons de iniciação (atg) e de terminação (tga) da transcrição indicadas. Os éxons que codificam o domínio *forkhead* (FOX) de ligação estão em destaque. A mutação relativa à família KE está localizada no éxon 14; o ponto de interrupção da translocação de CS está entre o éxon 3b e o éxon 4. Imagem (b) e texto extraídos e adaptados de LAI *et al.* (2001) e NEWBURY *et al.* (2002).

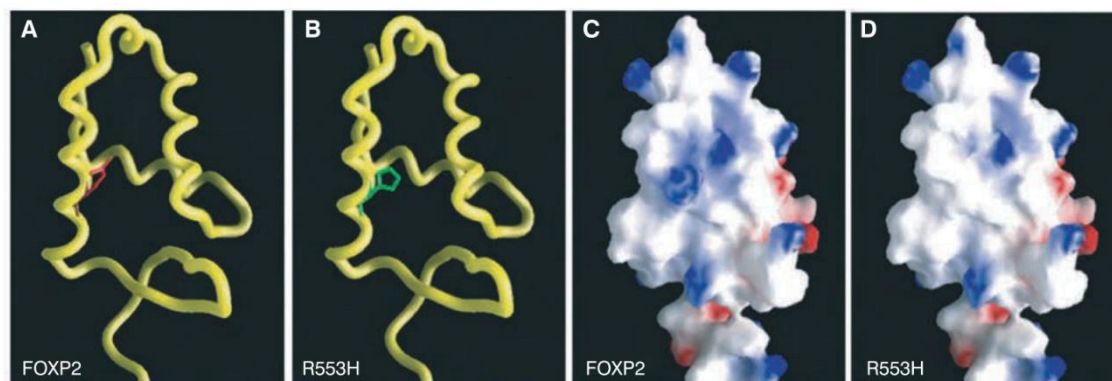


FIGURA 5.6. Modelos moleculares e representações da distribuição de carga eletrostática de superfície da proteína normal codificada pelo gene *FOXP2* (A e C) e da proteína resultante da mutação R553H (B e D) presente na família KE. Em (A) e (B) são mostrados dois modelos moleculares do domínio FOX da proteína, nos quais foram indicados os dois aminoácidos que podem aparecer na posição 553 e que fazem parte desse domínio: uma arginina (Arg), presente na variante normal (A) ou uma histidina (His), presente na variante mutada (B) encontrada na família KE. Em (C) e (D), a distribuição de carga eletrostática de superfície do domínio FOX da proteína normal (C) é comparada com a da variante que apresenta a mutação R553H (D). Essa mudança na configuração eletrostática do domínio FOX impede que a proteína se conecte ao DNA de outros genes. Extraído e adaptado de BANERJEE-BASU & BAXEVANIS (2004) e BURRACO (2012).

5.2. PAPÉIS BIOLÓGICOS DO *FOXP2*

Como já observamos, o *FOXP2* pertence a um grupo de genes (família FOX) que regulam a ação de outros genes através da produção de fatores de transcrição – proteínas que, como o nome sugere, ativam a transcrição (e, conseqüentemente, controlam a expressão) de outros genes. Os genes da família FOX estão presentes tanto em animais vertebrados quanto em invertebrados, mas o número de genes dessa classe parece aumentar ao longo da evolução (LEHMANN *et al.*, 2003). Essa família de genes se caracteriza por uma ampla variedade funcional, desempenhando diversos papéis durante a diferenciação celular, em especial durante a embriogênese. Como todas as células de um organismo possuem essencialmente o mesmo conjunto de genes, é preciso que um mecanismo controle quais devem ser expressos ou não, e em que momento. Um neurônio, por exemplo, tem uma demanda de proteínas completamente diferente daquela exigida por uma célula do fígado. Genes como o *FOXP2* permitem uma taxa diferencial de expressão gênica (i.e. de produção de proteínas) em cada célula, uma vez que regulam a síntese proteica de acordo com o metabolismo, a morfologia e a função de cada tipo celular. Como já vimos, taxas diferenciais de expressão gênica também

são as responsáveis pela diferenciação celular durante o desenvolvimento embrionário. A proteína FOXP2⁸⁰ (que tem o mesmo nome do gene que a produz) é expressa em altas concentrações no tecido neurológico durante o desenvolvimento do cérebro humano (o que, em tese, poderia estar relacionado com a formação de estruturas neurocognitivas responsáveis pela aquisição e pelo desenvolvimento da linguagem), mas também é expressa em células do coração, do intestino e do pulmão durante o desenvolvimento embrionário (SHU *et al.*, 2001), além de ser encontrada em outros órgãos do indivíduo adulto, onde atua como fator de transcrição de uma série de genes importantes para o funcionamento adequado de cada tipo celular específico.

Se a proteína por codificada pelo *FOXP2* é também importante para o desenvolvimento e o funcionamento de outros órgãos vitais, por que a disrupção do gene nos indivíduos afetados da família KE causa um distúrbio que se restringe “apenas” ao desenvolvimento da fala e da linguagem, provocando déficits profundos de articulação e problemas linguísticos? A resposta, segundo Lai *et al.* (2003), seria uma questão de dosagem. Indivíduos normais possuem duas versões “saúdáveis” (ativas) do gene *FOXP2*. Quando os genomas dos membros afetados da família KE foram analisados, descobriu-se que todos eles apresentavam uma mutação em apenas uma das versões do *FOXP2* (a outra continuava intacta), o que acarretava um déficit na produção da proteína reguladora, mas não a sua ausência completa. Assim, sugeriu-se que a quantidade de proteína expressa pelo único gene ativo seria suficiente para garantir o funcionamento adequado de certos tecidos, enquanto que, para um desenvolvimento cerebral normal (especialmente no que tange à formação e ao funcionamento de áreas relacionadas à linguagem), esse nível crítico mínimo de expressão não poderia ser alcançado com um único gene funcional. Alguns componentes biológicos de alta complexidade, como os relacionados ao cérebro, por exemplo, aparentemente exigem que a engenharia celular de expressão gênica trabalhe no nível máximo de produção, sintetizando doses elevadas de proteínas para que o componente se desenvolva normalmente. Essa hipótese é reforçada pela detecção de altos níveis de expressão do *FOXP2* no tecido cerebral de fetos humanos com desenvolvimento normal (FERLAND *et al.*, 2003; LAI *et al.*, 2003).

O *FOXP2* é expresso de forma particularmente abundante na fase intermediária do período gestacional (FERLAND *et al.*, 2003; TERAMITSU *et al.*, 2004), uma etapa em que

⁸⁰ Normalmente, uma proteína e o gene que a codifica são referidos pelo mesmo “nome”, diferindo entre si apenas pelo uso de caracteres em itálico (utilizados para se referir aos genes) ou sem itálico (utilizados para se referir às proteínas). Assim, por exemplo, o gene *FOXP2* codifica a proteína FOXP2.

há uma acentuação da migração neuronal, acompanhada de uma aceleração da diferenciação histológica e da regionalização do córtex, o que dá lugar à realização de processos neuroanatômicos centrais relacionados à organização estrutural e funcional do cérebro, como é o caso, por exemplo, do estabelecimento da assimetria anatômica e operacional existente entre os hemisférios cerebrais (SUN *et al.*, 2005). Os níveis mais altos de expressão gênica do *FOXP2* durante o desenvolvimento do cérebro humano na fase gestacional ocorrem nos núcleos da base (também chamados conjuntamente de gânglios basais) – um conjunto de agrupamentos neuronais (núcleo caudado; putâmen; globo pálido; núcleo subtalâmico; substância negra) localizados na base do cérebro (Fig. 5.7a), cujas principais funções estão relacionadas ao controle (inicialização e suavização) de movimentos musculares, supressão de movimentos involuntários, coordenação de mudanças de postura e integração de circuitos associados a tarefas cognitivas. A proteína FOXP2 é encontrada principalmente no putâmen (responsável pela regulação de movimentos amplos e envolvido na integração de circuitos relacionados à aprendizagem) e na cabeça e cauda do núcleo caudado (responsável pelo controle motor voluntário de movimentos e associado a circuitos cognitivos).

Em indivíduos que possuem uma versão defeituosa do gene *FOXP2*, o núcleo caudado se apresenta como uma região disfuncional que também exhibe anormalidades estruturais bilaterais (Fig. 5.8). A dispraxia orofacial manifestada por esses indivíduos pode estar correlacionada com o menor nível de atividade do corpo estriado (núcleo caudado + putâmen), um agrupamento de células que, entre outras tarefas, está associado ao planejamento de movimentos repetitivos e à execução de padrões motores estereotipados (CANALES & GRAYBIEL, 2000). Após o nascimento, a expressão gênica do *FOXP2* diminui ligeiramente com a idade em todas as estruturas/agrupamentos neuronais onde o RNA transcrito e a proteína são detectados durante o desenvolvimento embrionário. Até o momento, as análises realizadas em humanos parecem confirmar que o gene também é expresso no córtex cerebral, fundamentalmente nas regiões perisylvianas (vizinhas à fissura de Sylvius ou fissura lateral) de ambos os hemisférios (Fig. 5.7b; c), bem como nos lobos frontal e occipital (FERLAND *et al.*, 2003) e no cerebelo, região motora que atua como central de controle do movimento, possuindo extensas conexões com o cérebro e a medula espinhal (VARGHA-KHADEM *et al.*, 2005).

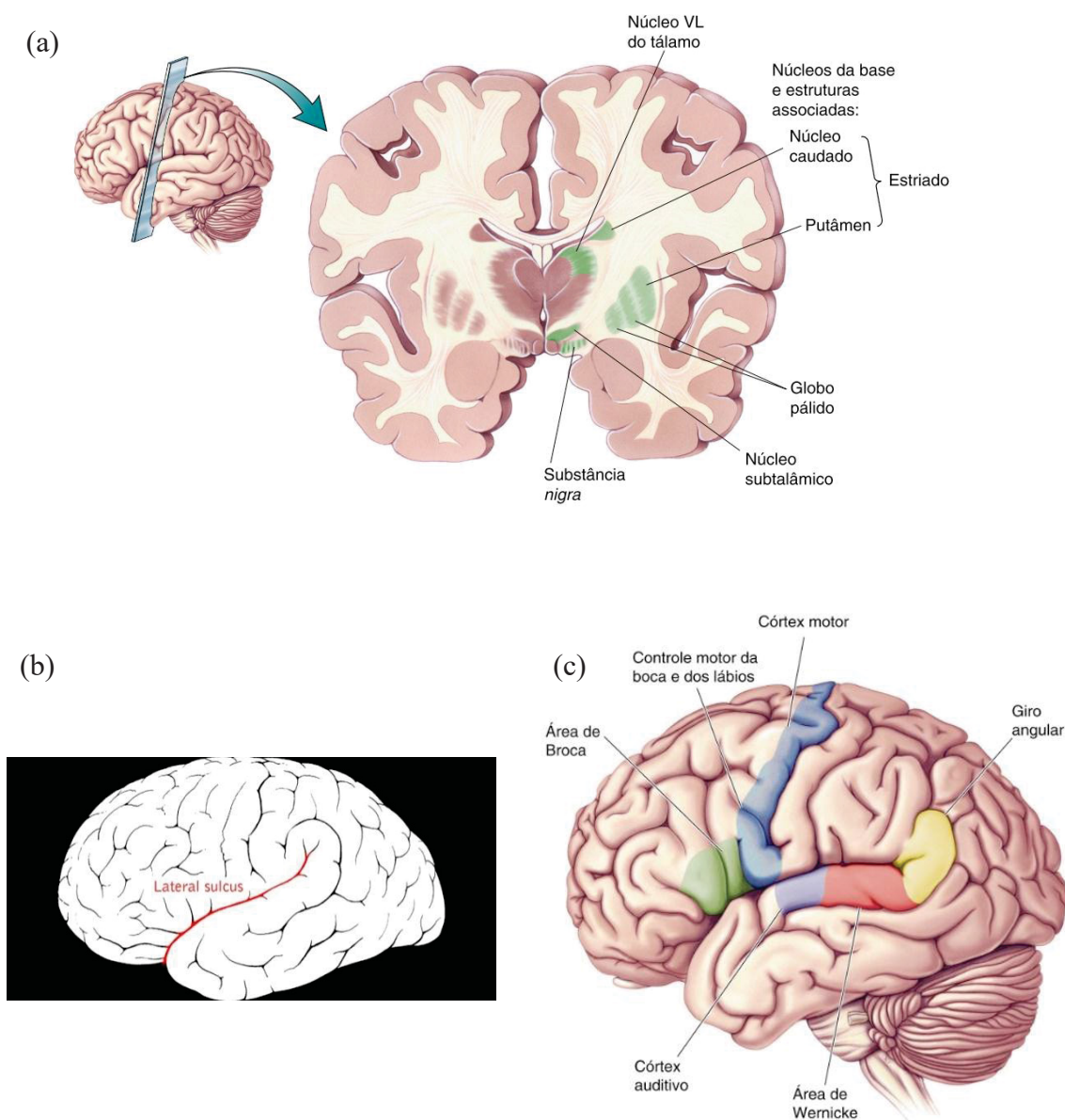


FIGURA 5.7. (a) Núcleos da base (gânglios basais). Os agrupamentos de neurônios dos núcleos da base têm como função principal controlar a seleção e a iniciação de movimentos voluntários. Feixes de neurônios conectam os núcleos da base às áreas corticais, mantendo-os permanentemente ligados e permitindo a integração do programa motor. No entanto, há fortes evidências de que esses núcleos não participam apenas de circuitos de função estritamente motora, mas também integram um grande número de circuitos neuronais paralelos, incluindo aí circuitos relacionados à memória e a outras funções cognitivas, como a linguagem. A expressão do *FOXP2* é particularmente elevada nessas áreas durante a formação do cérebro e os indivíduos com distúrbio de linguagem que possuem apenas uma cópia ativa do gene (como os da família KE) apresentam problemas estruturais e disfunções em alguns desses agrupamentos especiais de neurônios. **(b)** Fissura lateral. O *FOXP2* também é expresso de forma significativa em áreas corticais próximas à fissura lateral (ou fissura de Sylvius) durante determinadas fases do desenvolvimento do cérebro humano. Como pode ser observado em **(c)**, algumas dessas regiões perisylvianas são de extrema importância para o processamento neurológico da linguagem. **(c)**

Componentes-chave do sistema de linguagem no hemisfério esquerdo. No lobo frontal, a área de Broca – responsável principalmente pela produção/expressão da linguagem – encontra-se próxima à área que controla a boca e os lábios no córtex motor. A área de Wernicke – responsável sobretudo pela recepção/compreensão da linguagem – situa-se na superfície superior do lobo temporal, entre o córtex auditivo e o giro angular. Hoje, sabe-se que o processamento sensorial e motor da linguagem não ocorre apenas em regiões interconectadas do córtex cerebral (responsável pela execução de tarefas cognitivas de ordem superior) mas é realizado também em estruturas subcorticais que atuam na integração dos sinais (como os núcleos da base). Além disso, hoje sabemos que o processamento da linguagem não se dá através de um fluxo unidirecional de informações que vão sendo computadas ao longo de uma única via (cf. o Modelo de Wernicke-Geschwind do *Apêndice*), mas ocorre através de uma complexa rede de circuitos neuronais que funcionam de forma paralela e integrada (cf. o Modelo de vias paralelas do mesmo *Apêndice*). A linguagem, ao se realizar no cérebro, assemelha-se a uma descarga elétrica que se irradia em múltiplas direções, ativando inúmeras regiões. Extraído e adaptado de BEAR *et al.* (2017).

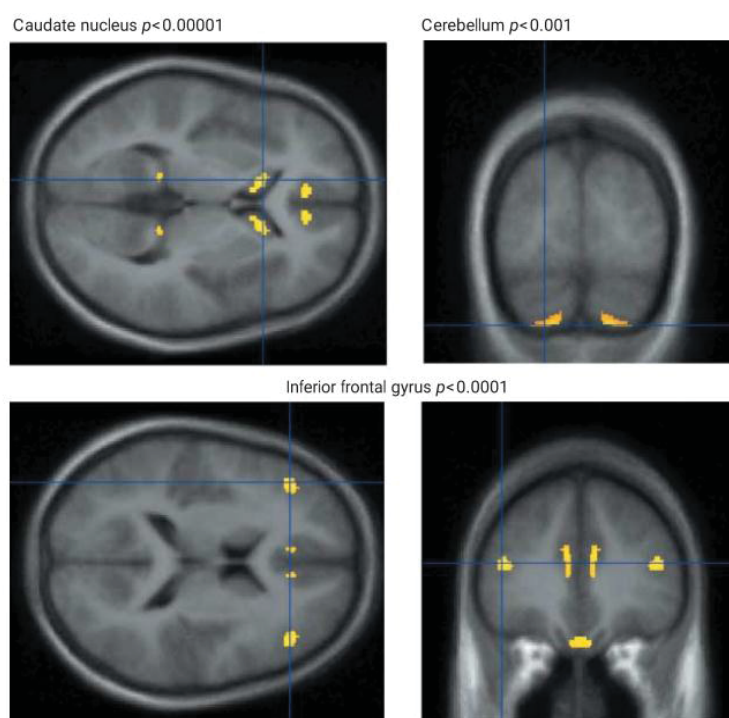


FIGURA 5.8. Em membros afetados da família KE, foi encontrada uma redução de substância cinzenta (i.e., de corpos de neurônios) no núcleo caudado (canto superior esquerdo), no cerebelo (canto superior direito) e na área de Broca no lobo frontal (parte inferior) – estruturas cujo desenvolvimento normal depende de níveis adequados de proteínas FOXP2 durante determinadas etapas do processo de formação do cérebro. Exames de neuroimagem feitos com os KE afetados antes mesmo da descoberta do *FOXP2* já haviam mostrado que eles apresentavam alterações morfológicas (reduções volumétricas) e funcionais em regiões específicas do cérebro. Durante a realização de uma determinada tarefa linguística, os membros afetados apresentavam uma subativação de áreas do córtex associadas à atividade motora, incluindo as áreas do hemisfério esquerdo (dominante com relação à linguagem) responsáveis pelo controle da boca e da porção inferior da face. Já a área de Broca, o córtex pré-motor esquerdo e o núcleo caudado – que fazem parte do circuito de produção da fala fluente – mostraram-se superativados nos KE afetados em relação aos indivíduos do grupo controle (VARGHA-KHADEM *et al.*, 1998). Imagem extraída de VARGHA-KHADEM *et al.* (2005).

5.3. ASPECTOS FILOGENÉTICOS DO *FOXP2*

Além de se expressar em tecidos que não têm relação alguma com a linguagem – em células do coração humano, por exemplo –, o *FOXP2* também faz parte do genoma e da história evolutiva de outras espécies, podendo ser encontrado, por exemplo, em crocodilos e beija-flores. A transversalidade filogenética do *FOXP2*, além de realçar a importância da função reguladora que este gene desempenha na codificação de fatores de transcrição responsáveis pela ativação e/ou supressão de outros genes, também pode oferecer informações valiosas sobre o seu papel na determinação da faculdade da linguagem na espécie humana.

Genes ortólogos são genes homólogos pertencentes a espécies diferentes que se originaram de um ancestral comum. O acúmulo de mutações ao longo da especiação acaba dando origem a versões alternativas de um mesmo gene, resultando em modificações nas proteínas correspondentes sintetizadas. Versões do *FOXP2* são encontradas em muitos animais (aves canoras, pombos, galinhas, camundongos, chimpanzés, bonobos, orangotangos), alguns deles utilizados como organismos-modelo em pesquisas genéticas. A sequência de aminoácidos e o padrão de expressão neural da proteína FOXP2 são extremamente conservados nos diferentes grupos, desde répteis até seres humanos, sugerindo um papel importante para o gene em cérebros de vertebrados, independentemente de estar relacionado ou não com a aprendizagem vocal (SCHARFF & HAESLER, 2005). No entanto, é interessante observar que, em aves canoras, o *Foxp2*⁸¹ é fortemente expresso em áreas do encéfalo envolvidas no aprendizado do canto (*vocal learning*) (Fig. 5.9). As aves canoras, assim como os humanos em determinadas fases da aquisição fonêmica, aprendem a vocalizar por imitação, utilizando regiões do cérebro intimamente ligadas à expressão do *Foxp2*, o que levanta questões importantes sobre as semelhanças e diferenças entre as diferentes versões do gene encontradas em espécies tão evolutivamente distantes. Embora mamíferos e aves tenham se separado de uma linhagem comum há mais de 300 milhões de anos, a proteína FoxP2 do diamante-mandarim (*Taeniopygia guttata*), por exemplo, difere da proteína Foxp2 encontrada em camundongos em apenas cinco posições de aminoácidos e da FOXP2 humana em oito posições; essas diferenças fornecem um valor médio de mais de 98% de identidade da proteína, mesmo entre aves canoras e humanos (HAESLER *et al.*, 2004). Além disso, o

⁸¹ Normalmente, genes identificados em seres humanos e em outras espécies (genes ortólogos) recebem o mesmo nome/código, diferindo-se entre si pelo uso de caracteres maiúsculos (seres humanos) ou minúsculos (outras espécies). Assim, por exemplo, o *FOXP2* é encontrado em humanos e o *Foxp2* é encontrado em canários.

padrão geral de expressão do *Foxp2* no cérebro do diamante-mandarim é notavelmente semelhante ao padrão de expressão detectado em cérebros de mamíferos, incluindo o cérebro do feto humano (TERAMITSU *et al.*, 2004). Existe a suspeita de que o *Foxp2* em algumas aves (e presumivelmente o *FOXP2* em humanos) esteja implicado na regulação da memória operacional, sendo responsável, ao menos em parte, pela constituição de um aspecto central de um *parser* (um analisador sintático), que seria utilizado por estas aves canoras para transmitir estruturas extremamente intrincadas através de ondas sonoras, que seriam reconstruídas no cérebro da ave receptora (PIATTELLI-PALMARINI & URIAGEREKA, 2011).

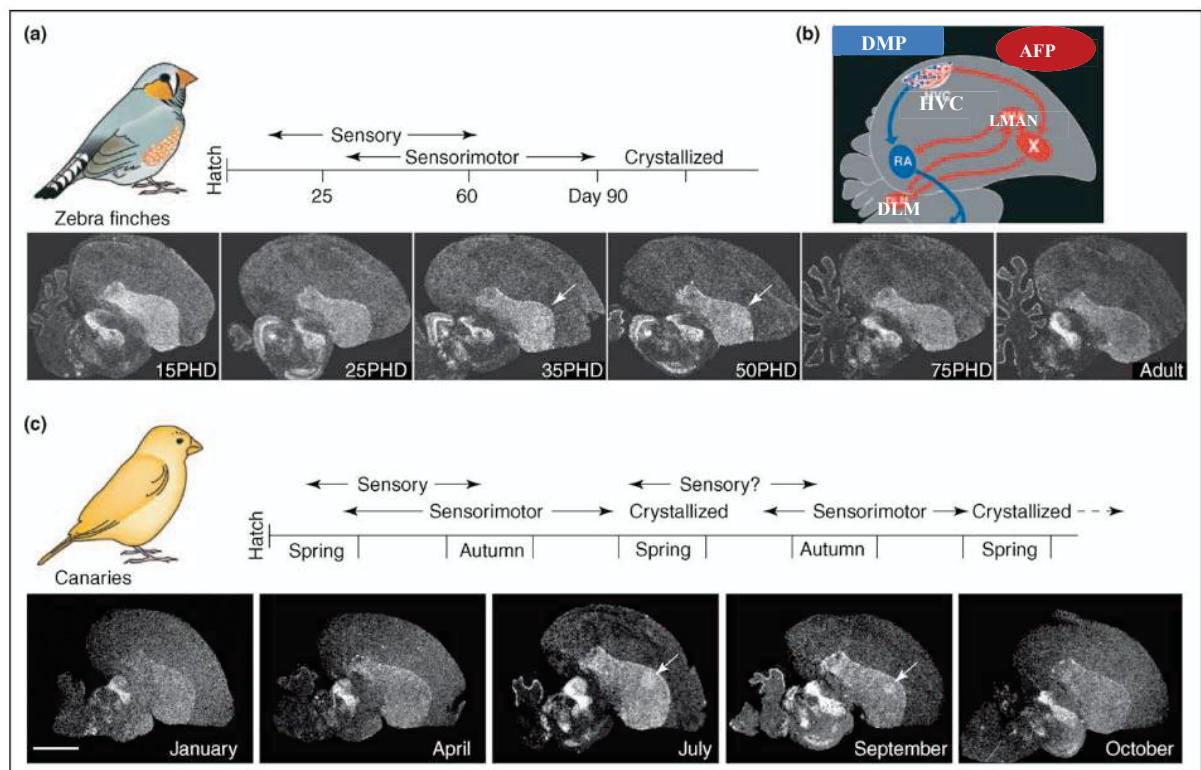


FIGURA 5.9. Aumento da expressão do gene *FoxP2* no corpo estriado de aves canoras durante períodos de plasticidade vocal. O processo de aprendizagem do canto de *Taeniopygia guttata* (a) ocorre a partir de uma ave adulta (tutor) e compreende três fases distintas: na fase de “aprendizagem sensível”, a ave memoriza o canto emitido pelo tutor, embora vocalize pouco; durante a fase “sensório-motora”, a ave começa a praticar o canto memorizado, a fim de eliminar possíveis desajustes em relação ao canto emitido pelo tutor; finalmente, ocorre uma fase de “cristalização”, na qual o canto se estabiliza, apresentando pouquíssimas variações na fase adulta. Esta curva de aprendizagem difere, porém, do que ocorre nos canários (c), que são capazes de incorporar novas “sílabas” ao seu canto ao longo da vida. Paralelamente, nos canários é observado um aumento sazonal da plasticidade neuronal dos circuitos responsáveis pela aprendizagem e/ou pela geração do canto (circuito do canto). Em ambos os casos, a seta branca indica a existência de altos níveis de expressão do gene *Foxp2* na área X do circuito do canto. (b) Anatomia e conectividade do circuito do canto. Os núcleos HVC e RA (*robust*

nucleus of arcopallium) fazem parte da via motora (DMP: *direct motor pathway*) necessária para a produção do canto (azul). HVC também fornece entrada para a via anterior do prosencéfalo (AFP: *anterior forebrain pathway*) (vermelho). A AFP compreende os núcleos LMAN (*lateral magnocellular nucleus of anterior nidopallium*), DLM (*dorsal lateral nucleus of the medial thalamus*) e área X. Este circuito é essencial para a aprendizagem do canto durante o desenvolvimento e para os períodos de plasticidade do canto na idade adulta. A expressão de *Foxp2* na Área X é elevada durante as fases de plasticidade do canto tanto em diamantes-mandarins juvenis (a) quanto em canários adultos (c). Imagem extraída de SCHARFF & HAESLER (2005). Texto adaptado de BURRACO (2012) e SCHARFF & HAESLER (2005).

Apenas dois aminoácidos distinguem a forma humana da proteína FOXP2 da versão encontrada em chimpanzés, gorilas e bonobos (STAES *et al.*, 2017). Mesmo que a trajetória evolutiva de seres humanos e chimpanzés os tenha separado há cerca de 6 milhões de anos, os estudos iniciais estimavam que as mutações diferenciadoras do gene *FOXP2* humano tivessem ocorrido há cerca de 200 mil anos (ENARD *et al.*, 2002), em um momento crucial de singularização do *homo sapiens* enquanto espécie. No entanto, pesquisas posteriores com o DNA de hominídeos arcaicos demonstraram que a sequência codificadora do *FOXP2* de neandertais é a mesma dos seres humanos modernos (KRAUSE *et al.*, 2007). Esses resultados sugeriram que, em vez de estarem sujeitas a uma seleção natural recente, as duas substituições de aminoácidos provavelmente já estavam fixadas no ancestral comum dos neandertais e dos humanos modernos, que viveu pelo menos há 400 mil anos (Fig. 5.10). O *Homo neanderthalensis*, ou simplesmente homem de neandertal, teria surgido há aproximadamente 400 mil anos e desapareceu há cerca de 40 mil anos (ROEBROEKS & SORESSI, 2016). Obviamente, o *FOXP2*, por si só, não pode ser considerado o responsável pela faculdade da linguagem ou pela capacidade expressiva; no entanto, suspeita-se que ele tenha desempenhado um papel importante no surgimento dessas aptidões no curso de nossa evolução (podendo, eventualmente, ser o responsável pelo ponto de virada que nos permitiu finalmente representar a linguagem em nossas mentes). Por isso, o fato de esse gene ser encontrado no cérebro de neandertais com a mesma sequência absoluta de nucleotídeos é, no mínimo, instigante. Apesar das consideráveis diferenças com relação à nossa espécie (mas levando-se em conta as consideráveis semelhanças), teriam estes hominídeos ancestrais – que coexistiram com o *homo sapiens* durante um longo período de tempo – a capacidade de falar?

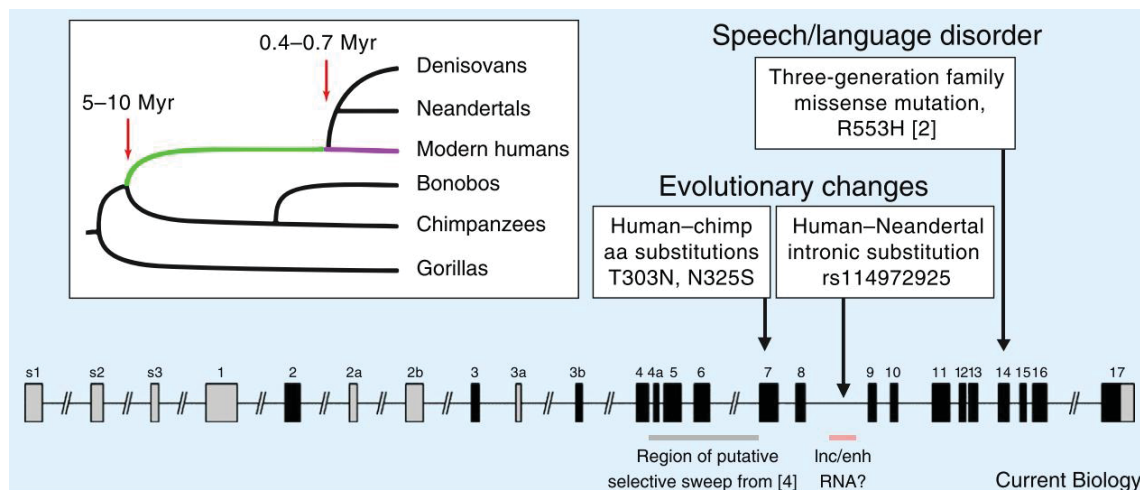


FIGURA 5.10. O *FOXP2* e a evolução humana. O locus do *FOXP2* humano abrange mais de 600 quilobases de DNA genômico. Os retângulos representam os éxons (presentes no RNAm processado), as linhas representam os íntrons (removidos do RNAm por *splicing*) e o sombreamento preto indica éxons que codificam trechos/regiões da proteína. Estão indicadas algumas mutações, inclusive aquela responsável pelo distúrbio de fala e linguagem identificado na família KE: R553H, uma mutação missense – i.e. responsável pela troca de um único nucleotídeo – no éxon 14, que teve como efeito a síntese de uma proteína disfuncional. As alterações evolutivas no *FOXP2* que foram hipotetizadas como relevantes para a origem da espécie humana estão destacadas, incluindo: duas substituições de aminoácidos (aa) no éxon 7, relacionadas ao contraste entre humanos e chimpanzés; um polimorfismo derivado intrônico entre os éxons 8 e 9, associado à distinção entre humanos modernos e neandertais; e um elemento conservado nesse íntron que pode representar um RNA longo não codificador (lnc) ou potenciador (enh) da atividade transcriptiva. Observe que as mutações evolutivas e as mutações causadoras do distúrbio são totalmente distintas umas das outras. O canto superior esquerdo mostra uma árvore evolutiva de homínídeos e parentes primatas mais próximos, com a linhagem que leva aos homínídeos arcaicos destacada em verde e a linhagem que leva exclusivamente aos humanos modernos em roxo. Myr: milhões de anos. Imagem e texto extraídos e adaptados de FISHER (2019).

Apesar de desconhecemos as reais capacidades linguísticas de nossos primos ancestrais, especula-se que pequenas e relativamente recentes mutações no gene *FOXP2*, associadas a todo um conjunto de mudanças neurológicas acumuladas evolutivamente, possam ter dotado a espécie humana de agrupamentos de neurônios (relacionados aos núcleos da base, por exemplo) capazes de controlar com precisão os movimentos musculares necessários à instanciação e à articulação dos sons da fala. Propostas mais audaciosas, no entanto, atribuem ao *FOXP2* funções que não se restringem apenas à organização mecânica da fala, mas que dizem respeito à especialização de regiões corticais responsáveis pelo desenvolvimento de competências cognitivas superiores relacionadas especificamente à linguagem.

Uma vez que uma das conclusões do estudo realizado por Liégeois *et al.* (2003) foi que a mutação do gene *FOXP2* foi associada à diminuição da atividade na área de Broca no hemisfério esquerdo (e em sua estrutura homóloga no hemisfério direito) durante tarefas de processamento linguístico, foi sugerido que uma maneira de explicar a trajetória evolutiva do gene na linhagem humana poderia ser postular que as mutações acumuladas ao longo dele teriam permitido que a nova proteína modificasse seu padrão de inibição/ativação durante o desenvolvimento de certas linhagens neuronais, especificamente, aquelas que dão origem à área de Broca, e assim amplificar as funções inicialmente associadas a esta região, facilitando assim o surgimento da sintaxe ou otimizando, pelo menos, o processamento fonológico e/ou o trabalho de memorização das palavras. Uma explicação alternativa poderia ser que essas mutações permitiram que a área de Broca fosse recrutada para a linguagem falada, consequentemente refinando a capacidade articulatória do indivíduo (CORBALLIS, 2004), o que, além do mais, estaria em consonância com a dispraxia orofacial observada naqueles que possuem a mutação (disfuncional) deste gene. Nesta linha de raciocínio, Enard *et al.* (2002) propuseram que a mutação do gene *FOXP2* teria sido selecionada porque possibilitaria um controle mais preciso dos movimentos orofaciais, facilitando, em última análise, o aparecimento do dispositivo articulatório que a espécie humana utiliza para a codificação e emissão de mensagens linguísticas. Por fim, uma hipótese adicional sobre o significado que a evolução do fator *FOXP2* pode ter tido na aparição filogenética da linguagem diz respeito à evidência de que ao longo da história evolutiva da espécie humana houve uma mudança significativa no controle da atividade motora responsável pela articulação dos sons, o que levou a uma remodelação da inervação motora do trato supralaríngeo, com o objetivo de transferir quase que completamente esse controle do sistema nervoso autônomo para o cérebro (LIEBERMAN, 2002). Este processo resultou, na prática, numa separação das emissões vocais de natureza linguística daquelas outras involuntárias, associadas às emoções despertadas em determinadas situações, que continuam a depender em certa medida do sistema nervoso autônomo, como ainda acontece com todas as vocalizações em primatas superiores (DEACON, 2000). A mutação recente do gene *FOXP2* poderia, portanto, estar relacionada à transferência do controle articular do sistema límbico para regiões corticais, e uma dessas regiões seria a área de Broca. (BURRACO, 2012: 324).

5.4. A FAMÍLIA KE

Para tentar definir quais características linguísticas podem estar relacionadas ao *FOXP2*, nada melhor do que investigá-lo a partir da família que serviu de base para a sua identificação.

Originalmente, o distúrbio apresentado pelos membros afetados da família KE foi classificado por Hurst *et al.* (1990) como uma forma severa de *apraxia verbal do*

desenvolvimento, uma vez que os indivíduos inicialmente investigados – um grupo de 6 crianças e adolescentes com idades entre 8 e 18 anos – apresentavam problemas articulatórios relativos ao posicionamento e à coordenação dos músculos do aparelho fonador e eram incapazes de programar e executar adequadamente os movimentos necessários à instanciação e à produção sequencial de fonemas, tornando a fala desordenada e ininteligível – o que era percebido, por exemplo, pela dificuldade que esses indivíduos demonstravam em produzir corretamente sequências complexas de sons, como encontros consonantais, que sofriam reduções, simplificações e substituições (*spoon* → “*boon*”; *blue* → “*bu*”). Além dessa “compactação” fonológica, os indivíduos afetados apresentavam uma tendência a omitir certos sons no início (*table* → “*able*”) e no final de palavras, o que aparentemente correspondia, neste último caso, a déficits relacionados ao sistema morfossintático de flexão e à morfologia derivativa. Em geral, as palavras polissilábicas eram transformadas em itens monossilábicos ou dissilábicos, indicando mais uma vez que os KE afetados lançavam mão de estratégias de redução fonêmica aplicadas a itens lexicais de maior complexidade sonora.

Havia uma lentificação na nomeação de fotos que retratavam objetos com os quais as crianças estavam familiarizadas e uma tendência a usar palavras do mesmo campo semântico em tarefas de nomeação (como, por exemplo, “*glass*” ou “*tea*” no lugar de *cup*, e “*sky*” para se referir a *star*). A compreensão também parecia estar afetada, principalmente com relação a estruturas comparativas (sentenças como “*The knife is longer than the pencil*” não foram compreendidas de forma suficiente) e passivas, nas quais a interpretação adequada dos papéis temáticos exigia uma computação sintática e semântica mais complexa (“*The girl is chased by the horse*” foi compreendida como “*The girl is chasing the horse*” e “*The boy chasing the horse is fat*” foi interpretada como sendo a mesma coisa que “*A thin boy chasing a fat horse*”). A maioria dos afetados era incapaz de reter/repetir expressões com mais de três palavras na sequência correta e não conseguia produzir construções gramaticais com mais de dois ou três itens lexicais combinados. Do ponto de vista sociointeracional, havia uma certa relutância em estabelecer conversas espontâneas. A autoconsciência dos KE afetados com relação às dificuldades linguísticas apresentadas aparentemente os tornava menos dispostos ao diálogo com desconhecidos. Todos os 16 indivíduos afetados da família apresentavam um atraso na aquisição e no desenvolvimento da linguagem, mas, excetuando-se os aspectos linguísticos, todos os outros marcos do desenvolvimento infantil (como engatinhar, sentar e caminhar) foram observados dentro dos limites de idade considerados normais para cada

etapa. Além disso, foram relatados poucos problemas neonatais e a audição e a inteligência de todos os membros afetados estavam dentro da faixa considerada normal.

Aos olhos de hoje, talvez o trabalho de Hurst *et al.* (1990) pareça relativamente modesto, dado o conjunto de pesquisas detalhadas que viriam nos anos seguintes, fazendo com que a família KE se tornasse um dos estudos de caso mais importantes da história da linguística – e de onde surgiriam as bases para a descoberta do *FOXP2*. A própria caracterização do distúrbio apenas como uma apraxia/dispraxia (os autores utilizam ambos os termos) verbal – restrita, portanto, à articulação dos movimentos da fala – parecia não corresponder clinicamente a determinadas situações concretas relatadas no próprio texto, que descrevem uma dificuldade aparentemente mais ampla dos indivíduos afetados em realizar combinações complexas de movimentos orofaciais, indicando a existência de uma dispraxia oral (i.e., não ligada exclusivamente ao controle da fala). Hoje, sabe-se que a dispraxia verbal dos KE é tributária de uma discinesia subjacente que compromete os movimentos de determinados grupos musculares da parte inferior do rosto. No entanto, apesar das limitações, as sementes de algumas ideias que viriam a ocupar o centro do debate quanto à natureza do fenótipo apresentado pelos KE – como os problemas relativos à morfossintaxe, por exemplo – já estavam presentes no trabalho pioneiro de Hurst *et al.* (1990), ainda que de forma difusa e pouco desenvolvida.

A existência de problemas que pareciam não derivar diretamente de limitações motoras – como a substituição lexical por aproximação semântica e a omissão de determinados morfemas finais – fez com que Gopnik & Crago (1991) propusessem uma investigação de natureza propriamente linguística para o transtorno apresentado pela família KE. A partir da análise dos dados obtidos em sua pesquisa, as autoras chegaram à conclusão de que o distúrbio não estava circunscrito apenas a um comprometimento do sistema de programação e de articulação mecânica da fala, mas atingia diretamente determinados componentes do sistema gramatical (em particular, o componente morfossintático), o que as levou a classificar o transtorno como uma disfasia do desenvolvimento (aparentemente condicionada por um único gene dominante, ainda não descoberto à época), rejeitando o diagnóstico inicial de apraxia proposto por Hurst *et al.* (1990). Definida como uma disfunção específica da produção e/ou da compreensão da fala e da linguagem, a disfasia do desenvolvimento é um transtorno linguístico que emerge na ausência de outras deficiências causais, como defeitos de audição ou em estruturas periféricas da fala, déficits intelectuais, subnormalidade mental, transtornos de personalidade, trauma cerebral, distúrbios

psicoafetivos ou transtornos psicóticos (TALLAL *et al.*, 1991). Com essa hipótese de trabalho, Gopnik & Crago (1991) traziam para o primeiro plano do debate a natureza linguística do distúrbio apresentado pelos KE, contrapondo-se a hipóteses alternativas que se fundamentavam em problemas de processamento cognitivo (distúrbios intelectuais), perceptual (distúrbios auditivos ou de recepção de sinais visuais) ou motor (apraxias ou dispraxias).

Publicado originalmente na *Cognition*, o artigo de Gopnik e Crago (1991) é o primeiro grande estudo de linguística sobre a família KE. Por isso, vamos analisá-lo mais detalhadamente nas próximas páginas.

Dos 30 indivíduos da família KE analisados à época (distribuídos ao longo de três gerações: avó, filhos e netos), 16 haviam sido diagnosticados como disfásicos ainda na idade escolar e tinham sido encaminhados para algum tipo de tratamento fonoaudiológico, com exceção da avó, que não recebeu nenhum tipo de auxílio especializado na infância. A análise de Gopnik & Crago (1991) contou com um conjunto de dados provenientes de 23 indivíduos pertencentes a este universo familiar, com idades entre 2 e 74 anos, sendo 15 deles disfásicos (incluindo uma bebê de 2 anos, cuja disfasia foi inferida por um atraso na aquisição da linguagem, e um casal de gêmeos fraternos composto por uma menina afetada e um menino não afetado). Segundo a avó, todos os seus filhos afetados (segunda geração) apresentavam na infância problemas de fala que não só eram semelhantes entre si, mas que também eram semelhantes àqueles observados entre os netos (terceira geração); além disso, os pais da segunda geração também relataram que seus filhos disfásicos apresentavam os mesmos problemas em determinadas fases da infância, sendo o mais saliente deles a ininteligibilidade da fala até aproximadamente os sete anos, quando então havia uma melhora da capacidade expressiva, apesar da existência de problemas fonêmicos que persistiam até a adolescência em alguns casos. Os problemas linguísticos dos membros afetados da família KE pareciam desaparecer ou eram quase imperceptíveis na idade adulta; no entanto, análises mais atentas demonstravam que essa normalidade era apenas aparente. Segundo Gopnik & Crago (1991), os KE afetados, por terem recebido uma educação especial desde muito cedo, aplicavam uma série de estratégias para contornar determinadas dificuldades linguísticas, mas o baixo desempenho em testes direcionados revelava que determinados componentes de sua gramática subjacente permaneciam severamente prejudicados. Se, do ponto de vista linguístico, os KE afetados conseguiam ao menos “camuflar” os problemas mais evidentes, do ponto de vista articulatório, no entanto, os déficits pareciam permanecer salientes na idade adulta, a ponto de

não haver discordância nos estudos sobre quais indivíduos da família eram afetados pelo distúrbio e quais não eram, indicando que o comprometimento motor da fala – aos quais Gopnik & Crago (1991) não dão muita importância – parecia ser mesmo um dos traços distintivos mais importantes para a delimitação adequada do fenótipo.

Como já salientamos anteriormente, genealogias como a dos KE (cf. Fig. 5.1), nas quais um determinado distúrbio pode ser observado em vários ramos diferentes de uma mesma família ao longo de várias gerações, fornecem indícios importantes para o estabelecimento de uma possível etiologia genética para o distúrbio analisado. No entanto, existem duas outras explicações possíveis para esses padrões familiares de recorrência: (i) o fenótipo analisado pode estar sendo causado por um fator ambiental ou (ii) o número aparentemente elevado de casos agregados pode ser o resultado de alguma anomalia estatística causada por uma amostra reduzida de indivíduos. Por isso, é preciso sempre levar em consideração as hipóteses (i) e (ii) no momento em que se avaliam os padrões familiares de um determinado traço fenotípico. Se vários membros da família investigada não são afetados pelo distúrbio e se não há nenhum fator ambiental que esteja atuando de forma diferencial sobre afetados e não afetados, então é mais provável que a causa seja genética e não ambiental (embora eventualmente exista a possibilidade de que o fenótipo esteja sendo condicionado por um componente ambiental ainda não identificado). Anomalias estatísticas, por sua vez, tendem a diminuir à medida que o tamanho da amostra (i.e., o número de indivíduos sob análise) aumenta. A composição familiar dos KE, cujos indivíduos disfásicos e não disfásicos haviam estado submetidos às mesmas condições ambientais (sociais e linguísticas) e cujo número de membros era relativamente elevado, colocava-os em uma posição privilegiada enquanto objeto de investigação genética.

O padrão de herança e a agregação familiar do distúrbio apresentado pela família KE, no entanto, representavam apenas algumas peças de um quebra-cabeça mais extenso. Para Gopnik & Crago (1991), a etiologia genética do transtorno só podia ser confirmada pela identificação precisa do fator molecular (i.e., do gene modificado) responsável pelo distúrbio. Além disso, era indispensável que o fenótipo linguístico fosse determinado com precisão, uma vez que, segundo as autoras, o déficit nuclear do transtorno parecia ser de natureza gramatical.

[...] Nem as evidências estatísticas de agregação familiar, nem os estudos de caso podem, por si só, provar a hipótese genética. É necessária uma evidência direta de que há uma diferença na estrutura genética que distingue os indivíduos normais dos disfásicos. Espera-se que esse marcador genético possa ser encontrado em breve, mas mesmo que tal marcador seja

identificado, estudos como este [Gopnik & Crago, 1991], que pretendem estabelecer a natureza exata do déficit linguístico que subjaz ao distúrbio, serão necessários para se ter uma ideia precisa do tipo de transtorno de desenvolvimento da linguagem que pode ser causado por essas diferenças na sequência de bases do DNA. Além disso, se for verdade que um fator genético autossômico dominante resulta em uma deficiência seletiva em uma parte específica da gramática, isso seria suporte para a hipótese de que a linguagem é feita de várias subpartes autônomas que podem ter evoluído em momentos diferentes. (GOPNIK & CRAGO, 1991: 9-10)⁸².

Os testes aplicados por Gopnik & Crago (1991) eram adaptações de testes desenvolvidos originalmente para a avaliação da competência linguística de indivíduos portadores de afasia adquirida. Todos os membros não afetados do grupo controle pertenciam à terceira geração e tinham menos de 17 anos. Apesar de contar com um conjunto mais extenso de dados provenientes de uma série de testes, as autoras se concentraram na caracterização linguística dos indivíduos mais velhos, cujas competências estariam mais próximas de um estado final da gramática, o que permitiria compreender melhor o(s) déficit(s) linguístico(s) subjacente(s). Assim, apenas os resultados referentes a um *sample* de seis indivíduos disfásicos (com idades entre 16 e 74 anos) e seis indivíduos controle (com idades entre 8 e 17 anos) foram reportados e analisados. Durante a apresentação dos resultados dos testes, as autoras tendiam a parear os sujeitos por ordem de idade. Assim, por exemplo, o disfásico mais jovem (16 anos) e o mais velho (74) eram emparelhados e comparados, respectivamente, com o não afetado mais jovem (8) e o não afetado mais velho (17). A diferença entre as idades dos dois grupos (a média de idade era de 39 anos entre os disfásicos e de 13 entre os não disfásicos), segundo as autoras, não prejudicaria a análise, já que todos, em tese, já teriam alcançado (ou estariam muito próximos de alcançar) o estágio final da gramática; além disso, por serem mais novos e não terem eventualmente alcançado a competência linguística dos adultos, seria mais provável que os indivíduos do grupo controle,

⁸² “[...] neither the statistical evidence of familial aggregation, nor the family studies in and of themselves, can *prove* the genetic hypothesis. What is needed is direct evidence that there is a characteristic difference in the genetic structure that distinguishes the normals from the dysphasics. It is to be hoped that a genetic marker will soon be found, but even if such a marker were found studies, like this one, of the exact nature of the underlying linguistic deficit would be necessary in order to have an idea of the precise impairment to language development that was caused by the observed differences in the genetic sequence of bases. Moreover, if it were true that an autosomal dominant genetic factor results in a selective impairment in a particular part of the grammar, this would be support for a hypothesis that language is made up of several autonomous subparts that may have evolved at different times.”

e não os do grupo disfásico, apresentassem resultados piores nos testes – o que não ocorreu, realçando ainda mais os déficits apresentados pelos KE afetados.

A primeira tarefa aplicada à família KE tinha o objetivo de avaliar a capacidade dos indivíduos testados de compreender e de discriminar as categorias de singular e de plural. O material utilizado no teste era composto por quatro tipos de objetos (livros, gizes de cera, moedas e balões), distribuídos em pilhas randômicas com um (singular) ou mais (plural) elementos do mesmo tipo. Na primeira parte do teste, os indivíduos eram solicitados a executar comandos simples (por exemplo, “*Please touch the books*”), que envolviam o processamento de apenas um nome marcado com o traço de plural ou singular; na segunda parte, os participantes precisavam obedecer a comandos mais complexos (como “*Put the crayon on the balloons*”), que exigiam a computação sequencial de dois itens lexicais com suas respectivas formas de singular ou plural. Desde Hust *et al.* (1990), havia uma suspeita de que os KE pudessem apresentar déficits morfológicos em seu sistema gramatical, o que eventualmente afetaria o processamento flexional de itens lexicais, prejudicando, por exemplo, a percepção e a produção do plural. Essa suspeita também dialogava com a sintomatologia de certos tipos de DEL, cujo quadro de déficits linguísticos se caracteriza por uma série de problemas de natureza fonológica e morfossintática, incluindo, por exemplo, dificuldades em compreender e utilizar verbos auxiliares e marcas morfológicas de tempo e de número. A ausência de marcas de plural e de passado regular na fala espontânea de crianças portadoras de DEL levou alguns pesquisadores a sugerirem a hipótese de uma falha no processamento auditivo de informações linguísticas como origem dos problemas gramaticais. A omissão de marcas morfológicas de flexão decorreria de uma falha no processamento perceptual de morfemas com pouca substância fonética. Assim, morfemas como o “-s” e o “-ed” do inglês, responsáveis, respectivamente, pela gramaticalização do plural e do passado regular, não seriam suficientemente salientes do ponto de vista sonoro, o que os impediria de serem processados – e adquiridos – pelos indivíduos com DEL.

No primeiro teste de discriminação do plural aplicado por Gopnik & Crago (1991), no entanto, tanto os indivíduos afetados quanto os não afetados apresentaram desempenhos considerados normais, com as médias dos resultados sendo praticamente as mesmas entre os dois grupos com relação à primeira parte da tarefa (processamento de comandos simples) e com uma ligeira vantagem do grupo controle em relação ao grupo disfásico na segunda parte, relativa ao processamento de comandos complexos. Em um primeiro momento, portanto, os membros disfásicos da família KE pareciam ser sensíveis à marcação de plural e eram

capazes de discriminar entre as formas marcadas com “-s” e as não marcadas, associando-as adequadamente a entidades singulares ou plurais no mundo. De toda forma, com os resultados obtidos nesse primeiro teste as autoras ao menos afastavam a hipótese de os problemas de linguagem apresentados pela família estarem sendo derivados de um déficit perceptual ou de processamento auditivo – i.e., de um déficit de natureza não propriamente linguística.

Uma vez confirmado que os indivíduos afetados da família KE eram capazes de reconhecer e de processar acusticamente traços de plural, Gopnik e Crago decidiram avaliar a sua capacidade de *formação* do plural. Para isso, as pesquisadoras aplicaram uma atividade que consistia em apresentar aos indivíduos testados figuras de animais imaginários que eram nomeados com pseudopalavras derivadas do inglês. Cada figura contendo um único animal (singular) era colocada ao lado de uma figura contendo duas ou três criaturas do mesmo tipo (plural). Para eliciar a pluralização das pseudopalavras, a pesquisadora apontava para a figura singular dizendo “*This is a zoop*”, e então apontava para a figura ao lado perguntando “*These are ____?*”. Foram utilizados seis pares de criaturas imaginárias, cujos nomes terminavam em consoantes desvozeadas (“zoop” → “zoop[s]” e “zat” → “zat[s]”), vozeadas (“wug” → “wug[z]” e “tob” → “tob[z]”) e sibilantes (“zash” → “zash[əz]” e “sas” → “sas[əz]”).⁸³

Como já observamos, pseudopalavras (*nonsense words*) são sequências sonoras construídas de acordo com as regras morfofonológicas de uma determinada língua, mas que não apresentam conteúdo semântico reconhecível no léxico da língua utilizada como base – ou seja, pseudopalavras são itens semanticamente vazios que podem ser pronunciados, repetidos, lidos, escritos e, principalmente, *transformados* de acordo com determinadas regras gerais. O uso de pseudopalavras em testes de processamento flexional é importante porque impede que os participantes recorram a formas flexionadas já previamente memorizadas e armazenadas em seu léxico mental; uma vez que esses itens nunca foram ouvidos antes, a construção de suas formas flexionadas só pode derivar da aplicação de algum tipo de regra computacional e não simplesmente de um acesso lexical a elementos memorizados.

⁸³ Representa-se entre colchetes [] a realização fonética de um fonema, em contraste com sua representação fonológica (entre barras oblíquas / /). Já os morfemas – as unidades mínimas de som e significado de uma língua – podem ser representados entre chaves { } ou sem nenhum símbolo que os flankeie (opção adotada por Gopnik & Crago (1991) e à qual também aderimos no presente trabalho). Nos exemplos utilizados no teste, [s], [z] e [əz] são instanciações fonéticas do mesmo morfema -s, responsável pela codificação gramatical da noção de plural no inglês.

Diferentemente do que havia sido observado por Gopnik & Crago no teste de reconhecimento de formas no plural, os resultados dos testes de pluralização de pseudopalavras mostraram uma diferença significativa entre os desempenhos de indivíduos disfásicos e não disfásicos. Apesar de compreenderem o conceito de plural e de serem capazes de manejar alguns itens corretamente, os KE afetados, aparentemente, não dispunham de regras internalizadas de construção do plural e seu desempenho dependia diretamente de um mecanismo de inferência que tentava mapear a regra utilizada em um determinado caso para depois aplicá-la indistintamente aos demais itens lexicais. Uma das participantes, por exemplo, ao conseguir produzir o plural adequado de “sas” (“sas[əz]”), acrescentou a terminação [əz] a todos os itens restantes: *zoo* foi transformado em *zoo*[əz], *tob* em *tob*[əz] e *zash* em *zash*[əz]. Uma outra participante disfásica, ao ser apresentada ao primeiro item, fez uma longa pausa enquanto repetia o enunciado em voz baixa para si mesma até sussurrar “*add an ‘s’*”. Ao pluralizar o item “sas”, a participante simplesmente prolongou a sibilante, produzindo “*sasss*”.

Segundo Gopnik & Crago (1991), o baixo desempenho dos indivíduos afetados nessa tarefa de pluralização seria fruto de um problema mais geral de processamento de traços gramaticais (de tempo, número e pessoa, por exemplo), constituindo uma evidência do que as autoras chamaram de *feature-blindness grammar* – uma hipótese de trabalho segundo a qual o distúrbio apresentado pelos KE possuiria como uma de suas características principais a ausência de regras internalizadas capazes de analisar e interpretar determinados traços gramaticais acoplados aos itens lexicais, que seriam armazenados no léxico como itens “indecomponíveis”. Assim, formas marcadas com o morfema “-s”, por exemplo, eram processadas por esses indivíduos disfásicos como itens lexicais não analisáveis que incluíam a noção de plural codificada diretamente em seu significado. “[Uma forma como] ‘books’, portanto, é interpretada por esses sujeitos como um item lexical indecomponível que significa ‘vários objetos legíveis’ e não como um item que é gerado por uma regra morfológica [do tipo Nome + -s] a partir de uma representação subjacente da palavra ‘book’ e do morfema codificador de plural” (GOPNIK & CRAGO, 1991: 36)⁸⁴.

Esse déficit de leitura de traços também pôde ser observado em tarefas de morfologia derivacional e de flexão verbal, nas quais os indivíduos disfásicos apresentaram resultados

⁸⁴ “‘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*.”

significativamente inferiores aos dos indivíduos do grupo controle. No teste de morfologia derivacional, os participantes eram expostos a um conjunto de duas sentenças logicamente inter-relacionadas e precisavam completar a segunda a partir de um item lexical presente na primeira, que servia como base formal ou conceitual para a derivação. Algumas sentenças envolviam regras derivacionais (“*There is a lot of sun. It is very ____.*” [sunny]) e outras envolviam algum tipo de acesso lexical (“*He has lots of wisdom. He is very ____.*” [wise]). A média de acertos do grupo disfásico foi de 2 na primeira pontuação⁸⁵ apresentada pelas pesquisadoras (incluindo dois indivíduos que não produziram nenhuma das respostas esperadas) e a média do grupo controle foi de 6,167; na segunda pontuação, a média dos disfásicos foi de 3 e a do grupo de não afetados foi de 7,17. Nas duas situações, a pontuação máxima possível era 10. Frequentemente, os indivíduos disfásicos lançavam mão de estratégias lexicais para responder àqueles testes que, na realidade, exigiam a geração derivacional de itens morfologicamente relacionados ao estímulo. Ao que parece, o déficit de traços proposto por Gopnik & Crago (1991) poderia se estender ao sistema de processamento derivacional dos indivíduos afetados, já que eles apresentavam dificuldades mais gerais de computação morfológica. Muitas vezes, ao serem elicitados, os disfásicos davam respostas que eram advindas de um processamento lógico-semântico – e não morfológico – das sentenças. Nestes casos, as respostas, mesmo quando eram contextualmente apropriadas (“*You can wash this. It is dirty*” [washable]), foram consideradas inadequadas.

Contrastes expressivos entre o grupo de disfásicos e o de não disfásicos também foram verificados em tarefas de flexão verbal. Foram aplicados testes que envolviam mudanças de tempo (e eventualmente de aspecto) verbal. A partir de uma forma verbal dada na primeira sentença, os indivíduos eram elicitados a completar lexicalmente a segunda de duas sentenças relacionadas. Em alguns casos, eram necessários ajustes na forma verbal a partir de regras de computação do tipo Verbo + “-ed”, responsável pela formação do passado regular (“*Every day he walks eight miles. Yesterday he ____.*” [walked]); em outros, a escolha da forma adequada dependia de um processo de seleção lexical que era acionada e orientada por um determinado verbo irregular (*go* → *went*; *is* → *was*). Os indivíduos também foram solicitados a produzir construções verbais relacionadas ao presente, ao futuro (“*Yesterday the girl baked a cake. Tomorrow she ____.*” [will bake a cake]) e ao presente contínuo (aspecto progressivo).

⁸⁵ Em alguns testes, Gopnik & Crago (1991) adotaram dois conjuntos distintos de critérios para avaliar os resultados, produzindo duas pontuações diferentes (dois *scores*).

A pontuação média dos indivíduos não afetados foi de 7,83 no primeiro *score* e de 9,18 no segundo; já a pontuação média dos indivíduos afetados foi de 3,16 e 3,83 no primeiro e no segundo *score*, respectivamente (pontuação máxima possível = 10). As diferenças entre as pontuações do grupo controle e do grupo disfásico com relação à computação adequada de marcadores verbais de tempo realçavam mais uma vez o déficit morfológico dos KE afetados. As estratégias usadas por esses indivíduos durante a realização dos testes de flexão verbal – uso recorrente de formas não marcadas do verbo utilizado como estímulo no teste (algo que não ocorreu em nenhum momento entre os indivíduos controle) e/ou substituição do verbo de base por um verbo semanticamente relacionado (algo que ocorreu com uma frequência bem maior do que aquela observada entre os KE não afetados) – eram semelhantes às aquelas observadas nos testes de morfologia derivacional (inclusive com relação aos tipos de erros gerados por sua aplicação) e pareciam indicar um déficit processual subjacente que atravessava tanto o sistema morfológico de flexão quanto o de derivação.

Em ambos os testes, os disfásicos pareciam não ser capazes de entender que o objetivo do teste era manipular uma regra gramatical implícita, embora os normais entendessem sem hesitação. Mesmo quando solicitados com exemplos e instruções específicas, eles pareciam incapazes de perceber que havia um padrão subjacente que era governado por uma regra geral. (GOPNIK & CRAGO, 1991: 44)⁸⁶.

Foram testadas também as capacidades de compreensão e de processamento de comandos sequenciais complexos. O teste consistia em executar adequadamente comandos do tipo “*Here are three crayons. Drop the yellow one on the floor, give me the blue one and pick up the red one.*” O participante era instruído a esperar até que todo o comando fosse dado. Assim, a execução da tarefa dependia também da capacidade de armazenamento da memória de trabalho. Não foram observadas diferenças significativas entre os desempenhos dos indivíduos controle e dos indivíduos afetados e em quase todas as situações a sequência de movimentos foi executada perfeitamente. Esses resultados apontavam mais uma vez para a existência de uma assimetria de competências linguísticas nos indivíduos disfásicos, com a preservação das capacidades de recepção e compreensão e o comprometimento das capacidades de geração e expressão.

⁸⁶ “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.”

A preservação da capacidade de recepção e compreensão dos KE afetados foi atestada também em tarefas de interpretação sintática. Os participantes recebiam uma folha com ilustrações e eram instruídos a indicar qual delas representava a situação descrita na sentença enunciada por uma das examinadoras. Em um dos testes, por exemplo, o participante precisava interpretar construções com pronomes reflexivos (“*He_i washes himself_i*”) ou com pronomes não reflexivos (“*He_i washes him_j*”), indicando qual das figuras recebidas – que mostravam (i) um homem lavando a si mesmo, (ii) uma mulher lavando a si mesma, (iii) um homem lavando um menino e (iv) uma mulher lavando uma menina – correspondia adequadamente ao significado da sentença. A interpretação sintática também foi testada com construções que contrastavam pronomes quanto aos traços de gênero e número (“*He holds him*”; “*He holds her*”; “*He holds them*”); construções ativas e passivas na forma negativa (“*The truck does not pull the car*” e “*The truck is not pulled by the car*”) e construções/comandos que exigiam o processamento de pronomes possessivos (“*Show me the mother’s baby*” e “*Show me the baby’s mother*”). Não houve, em nenhuma das modalidades testadas, diferenças significativas de desempenho entre o grupo normal e o grupo disfásico. Aparentemente, o mapeamento sintático e o processamento semântico de sentenças estavam preservados nos indivíduos portadores do distúrbio. Esse achado, quando confrontado com a brevidade dos enunciados dos KE afetados (que, como foi observado inicialmente por Hurst *et al.*, 1990, raramente produziam expressões com mais de três palavras), talvez indicasse que a extensão reduzida de suas sentenças não derivava necessariamente de uma limitação do sistema sintático (combinatório), mas era fruto de um déficit do sistema expressivo que impedia os indivíduos afetados de instanciar corretamente as construções geradas em um nível mais abstrato. Outro ponto a se considerar é que os KE tinham consciência de suas limitações e muitas vezes optavam por expressões curtas a fim de evitar situações de constrangimento social.

Os participantes também foram submetidos a testes de julgamento de (a)gramaticalidade, em que deveriam dizer se determinadas construções da língua inglesa estavam corretas ou não – e corrigi-las, quando fosse o caso. Foram apresentados 30 itens, sendo 9 deles considerados corretos (gramaticais) e 21, incorretos (agramaticais). As construções irregulares continham erros de número (“*The boy eats three cookie*” [cookies]); de concordância verbal (“*The boy kiss the pretty girl*” [kisses]); de tempo (“*Yesterday the girl pet a dog*” [petted]) e de aspecto (“*The little girl is play with her doll*” [playing]). No primeiro *score*, que levou em consideração apenas os julgamentos de gramaticalidade (pontuação

máxima = 30), a pontuação média do grupo disfásico (17,17) foi significativamente inferior à pontuação média do grupo controle (27,5); no segundo *score*, que considerava as respostas relativas às correções das construções agramaticais (pontuação máxima = 21), também foi observada uma diferença expressiva entre a pontuação do grupo disfásico (média de 7,83) e a do grupo controle (média de 18,33). Os disfásicos tendiam a modificar mais de uma parte da construção em suas tentativas de gramaticalização e as pesquisadoras tentaram adotar o critério mais generoso em sua análise; mesmo assim, há uma clara assimetria de desempenhos. Esses resultados pareciam corroborar a hipótese de trabalho das autoras (*feature-blindness* ou *feature-deficit grammar*), uma vez que os indivíduos afetados aparentavam não conseguirem processar adequadamente os traços funcionais (de número, pessoa, tempo e aspecto) que tornavam as construções gramaticais ou agramaticais – e muito menos conseguiram manipulá-los quando eram necessárias correções. Apesar de serem sensíveis à substância fonética de tais traços, os KE disfásicos eram insensíveis à função morfossintática codificada neles.

Em um outro teste de gramaticalidade, relativo à interpretação de papéis temáticos, os participantes precisavam dizer se construções como “*The girl eats a cookie to the boy*” e “*The boy puts the book*” eram corretas (gramaticais) ou incorretas (agramaticais ou anômalas), corrigindo a sua estrutura argumental quando necessário. Tanto os indivíduos disfásicos quanto os indivíduos do grupo controle lançaram mão de duas estratégias principais para ajustar as relações temáticas das construções anômalas: fornecer os argumentos que faltavam (“*The nice girl gives*” → *The nice girl gives him a ball*”) ou substituir o núcleo verbal da construção por um verbo que fosse adequadamente saturado pelos argumentos disponíveis (“*The girl eats a cookie to the boy*” → *The girl gives a cookie to the boy*). Os indivíduos afetados e os indivíduos normais tiveram desempenhos satisfatórios e muito semelhantes entre si, indicando que o déficit de linguagem dos KE não afetava áreas responsáveis pela estrutura argumental e pela distribuição de papéis temáticos na sentença.

Segundo as autoras, os resultados combinados dos dois testes de gramaticalidade – desempenho abaixo do esperado no julgamento de erros relacionados a traços morfossintáticos e desempenho adequado no julgamento de erros relacionados ao processamento temático – sugeriam que os KE disfásicos eram capazes de processar erros em estruturas argumentais, mas não eram capazes de reconhecer erros na marcação gramatical de traços funcionais (em especial, traços de número, pessoa, tempo e aspecto).

Gopnik & Crago (1991) também avaliaram a competência narrativa da família KE. Primeiro, foram aplicados testes que tinham o objetivo de avaliar a capacidade dos indivíduos de compreender narrativas breves. Para isso, os participantes precisavam ouvir atentamente a história narrada para logo em seguida responder a algumas perguntas simples relacionadas a acontecimentos explicitamente mencionados na narrativa. Ambos os grupos – disfásico e controle – tiveram desempenhos considerados normais. Na segunda tarefa, os participantes precisavam contar uma história curta a partir de um conjunto de seis imagens que representavam uma sequência de eventos conectados (Fig. 5.11). Um dos aspectos avaliados pelas autoras estava relacionado à anáfora (compreendida aqui como o processo de recuperação de informação lexical em uma estrutura). Assim, foram contabilizadas as ocorrências de sintagmas nominais completos (*full noun phrases*, como “*the boy*”) e sintagmas nominais pronominais (como “*he*”). Os resultados mostraram que os disfásicos faziam um uso significativamente maior de sintagmas nominais completos quando comparados aos indivíduos não disfásicos, que optavam preferencialmente pelos pronomes nos contextos autorizados. Nos exemplos a seguir, sublinhamos os sintagmas pronominais.

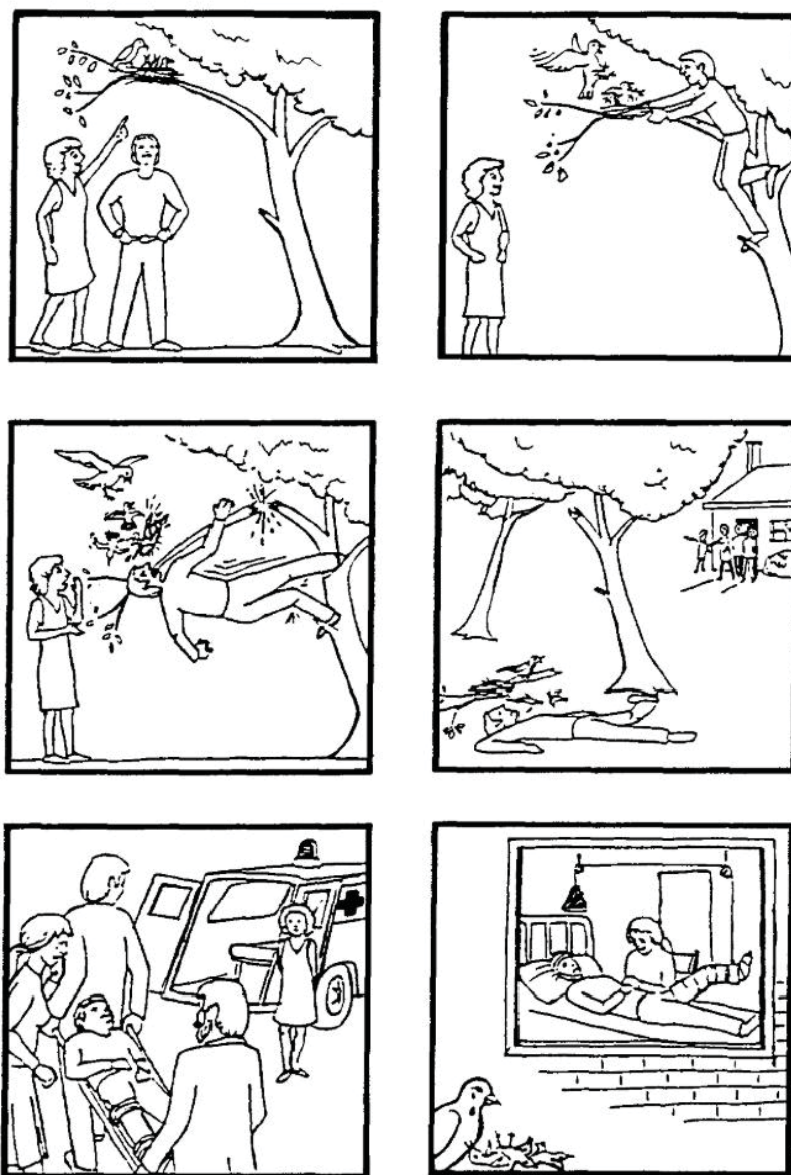


FIGURA 5.11. Sequência de imagens utilizada como estímulo para a produção de narrativas curtas. Extraído de Gopnik & Crago (1991): 42.

Exemplo de narrativa curta produzida por um integrante disfásico.

The lady's pointing at the bird and the man is watching her. 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 he's got a broken leg in the hospital. (GOPNIK & CRAGO, 1991: 43).

Exemplo de narrativa curta produzida por um integrante do grupo controle.

The woman's looking at the birds and she's saying to him. "Oh, they're lovely birds, but they got stuck in the tree." So he says, "Well. I'll get them down then." So he climbs up the tree and the branch breaks. So she's nearly in tears 'cause she swallowed the wrapping string, and the birds fly off and they fall to the ground 'cause they can't fly. So he falls on the ground and he breaks his leg. So she runs indoors to call an ambulance with all the neighbors outside the door. The ambulance comes. They get in the van and he's in the hospital, and the birds crying by the window. (GOPNIK & CRAGO, 1991: 44).

O uso recorrente de NPs completos e, principalmente, a baixa frequência de NPs pronominais eram indícios que reforçavam a ideia de que os indivíduos disfásicos apresentavam um déficit de processamento de traços que os impedia de interpretar e usar adequadamente determinadas entidades gramaticais, como os pronomes. Segundo Gopnik & Crago (1991), os KE afetados eram capazes de processar e interpretar apenas a dimensão semântica dos pronomes, mas não os seus traços funcionais. Assim, um elemento como “*he*”, por exemplo, era lexicalizado como “um indivíduo do sexo masculino não especificado”, sem que os seus traços dêiticos e referenciais pudessem ser acessados e lidos pelo sistema gramatical. Por isso, em tarefas mais simples, em que era preciso, por exemplo, indicar corretamente qual das figuras (um homem ou uma mulher) correspondia a um determinado pronome (“*he*”), os indivíduos afetados apresentavam desempenhos adequados, já que a resolução do teste dependia apenas do significado codificado no pronome lexicalizado. Em uma relação anafórica, entretanto, a interpretação referencial adequada de um pronome depende de uma computação mais complexa de marcadores gramaticais, já que é preciso haver uma compatibilidade de traços (*feature match*) entre o pronome e o sintagma nominal lexical que lhe serve de referência. Assim, o déficit de marcadores gramaticais codificados nos pronomes fazia com que os indivíduos disfásicos não conseguissem “emparelhar” os pronomes com outros sintagmas nominais, bloqueando a correferenciação e, conseqüentemente, a anáfora.

Adicionalmente a esses testes, Gopnik & Crago mediram o tempo de resposta dos participantes em quatro das tarefas mencionadas: julgamento de gramaticalidade, morfologia derivacional, flexão verbal e julgamento de gramaticalidade com relação à estrutura argumental. Com exceção desta última tarefa, cuja resolução dependia principalmente de competências sintáticas e semânticas, em todas as outras a duração média das respostas dos

indivíduos disfásicos foi notadamente maior do que a observada entre os indivíduos controle, indicando que os KE afetados precisavam de muito mais tempo para processar respostas referentes a problemas que exigiam a interpretação de marcadores gramaticais e a manipulação de regras – e mesmo assim, suas soluções se mostravam muitas vezes inadequadas.

Gopnik & Crago (1991) também avaliaram a escrita de duas crianças disfásicas da família KE – uma menina de 11 anos (identificada por C) e um menino de 10 (T) – que haviam recebido cadernos nos quais precisavam escrever uma história por semana durante todo o ano escolar. Um dos parâmetros avaliados pelas autoras foi a forma como as crianças utilizavam os verbos, em conjunto com as correções feitas pelo professor. Os dados mostravam que a taxa de verbos irregulares usados de forma correta pelas duas crianças era significativamente mais alta do que a taxa de acertos relativa aos verbos regulares. Além disso, a porcentagem de verbos irregulares utilizados corretamente logo na primeira ocorrência (antes da correção e da orientação do professor) também foi expressivamente mais alta do que o observado entre os verbos regulares. Essa assimetria se manteve em todos os recortes propostos pelas autoras.

Segundo Gopnik & Crago (1991), essa distribuição assimétrica do uso correto de formas verbais regulares e irregulares na escrita seria um reflexo do déficit processual dos indivíduos disfásicos da família KE. Mesmo quando as formas regulares corretas eram aprendidas em resposta às correções específicas feitas pelo professor, a reincidência sistemática de erros semelhantes em outros verbos regulares indicava que a criança tinha apenas armazenado aquelas formas lexicais corrigidas, mas não havia adquirido a regra geral subjacente. Já o processamento dos verbos irregulares, por supostamente não depender da aplicação de regras de afixação de marcadores gramaticais, mas sim de princípios de seleção lexical e de recuperação de palavras do dicionário mental (*eat* → *ate*), seria realizado adequadamente pelas crianças afetadas.

A análise integrada de todo esse conjunto de dados da família KE fez com que Gopnik & Crago (1991) rejeitassem a hipótese de um déficit derivado de limitações exclusivamente motoras e/ou perceptuais. Como demonstrado pelos testes, os indivíduos disfásicos eram sensíveis às formas fonológicas dos itens linguísticos. Eles podiam ouvir e produzir palavras como “*books*” e “*walked*”. O que eles não podiam fazer, segundo as autoras, era representar esses itens abstratamente com todos os seus traços gramaticais processados e analisados.

Assim, itens lexicais gramaticalmente complexos (i.e., ricos em traços) eram simplesmente armazenados como itens lexicais gramaticalmente indecomponíveis (i.e., não analisáveis). Esse déficit de varredura e de interpretação de traços (*feature-deficit*) impedia os indivíduos afetados de manipular adequadamente as representações abstratas dos itens lexicais “absorvidos”, dificultando (ou mesmo anulando) a construção de regras e de paradigmas mentais capazes de processar morfologicamente novos itens – o que podia ser facilmente observado pelo baixo desempenho dos probandos em tarefas de morfologia flexional e derivacional. Por não conseguirem processar as regularidades subjacentes a determinadas construções da língua (no caso, o inglês), os indivíduos disfásicos da família KE nunca chegaram a construir uma gramática mental suficientemente operacional com relação a determinados componentes – uma gramática capaz de compreender que cada instanciação de uma língua – cada construção linguística – é uma representação da regra abstrata que a gerou.

Por fim, Gopnik & Crago (1991) procuram integrar a insensibilidade aos traços (*feature-blindness*) do sistema morfossintático dos KE à teoria de aquisição de Pinker (1984), baseada na construção de paradigmas cuja formatação depende de restrições gramaticais inatas.

Pinker (1984) argumenta que os dados referentes à forma como a flexão é normalmente adquirida requerem uma teoria de aquisição que postule que as crianças são construtoras de paradigmas. O modelo especifica restrições detalhadas que orientam o modo como tais paradigmas são construídos e como eles são estendidos. Mas o que é fundamental para essa hipótese é que a criança deve saber que a linguagem provavelmente tem regularidades subjacentes e que essas regularidades são representáveis na forma de paradigmas. A criança deve saber ainda que esses paradigmas não são simplesmente generalizações sobre casos observados, mas que representam propriedades gerais do sistema morfológico que podem ser preditivas de novos casos. [...] Os dados apresentados acima se mostraram consistentes com a proposta de uma explicação predominantemente linguística para uma possível forma do transtorno apresentado pela família. Essa análise sugere que os disfásicos não possuem o mecanismo normal de aquisição de linguagem descrito por Pinker (1984), que lhes permitiria (ou até os obrigaria a) construir paradigmas flexionais com base em regularidades hipotetizadas a partir da evidência linguística observada. A incapacidade de construir tais paradigmas resultaria em um comprometimento seletivo daquele componente da gramática que codifica a morfologia abstrata, poupando outras habilidades linguísticas, como aquelas relacionadas à determinação das estruturas de argumentos. (GOPNIK & CRAGO, 1991: 46-47)⁸⁷.

⁸⁷ “Pinker argues that the data from the way in which inflection is normally acquired require a learning theory that postulates that children are paradigm builders. The model specifies detailed constraints on the way in which

O estudo de Gopnik & Crago (1991) representou um ponto de virada na forma como o transtorno apresentado pelos KE vinha sendo caracterizado até então. De fato, as autoras deixam de lado o diagnóstico de dispraxia verbal proposto inicialmente por Hurst *et al.* (1990), que se baseava principalmente em um quadro de déficits motores relacionados à produção da fala, e apresentam de modo pioneiro uma hipótese propriamente linguística, mudando o diagnóstico do distúrbio para um tipo especial de disfasia. Além disso, baseadas em uma concepção gerativista de linguagem, Gopnik & Crago (1991) propõem que os déficits linguísticos dos indivíduos disfásicos da família KE derivam de um problema no sistema computacional da gramática, que é incapaz de processar eficientemente o *input* lexical de modo a extrair dele diretrizes para a construção de paradigmas, afetando principalmente a morfologia flexional e a derivação.

Essa caracterização *linguística* do distúrbio apresentado pelos KE se baseava, em última análise, em dois pressupostos teóricos fundamentais: o primeiro dizia respeito à ideia de modularidade da mente, segundo a qual o sistema cognitivo humano corresponde a um conjunto de módulos especializados, sendo um deles o responsável pela realização da linguagem no cérebro. Apesar de estabelecer uma série de interfaces com outros sistemas – como o sistema sensorio-motor, responsável pela programação e pela execução de movimentos musculares –, esse módulo linguístico é relativamente autônomo e possui sua própria natureza. Frente a este modelo de arquitetura da mente, as dispraxias passam a ser vistas como distúrbios derivados de problemas no sistema sensorio-motor (extralinguístico); o que Gopnik & Crago (1991) fazem em sua proposta é transferir o déficit nuclear do distúrbio apresentado pela família KE para o domínio cognitivo da linguagem. O segundo pressuposto teórico que fundamentava a análise linguística das autoras era, de certa forma, uma extensão

such paradigms are constructed and the way in which they are extended. But what is fundamental to this hypothesis is that the child must know that language is likely to have underlying regularities and that these regularities are representable in the form of paradigms. The child must know further that these paradigms are not simply generalizations over observed cases, but that they represent general properties of the morphological system that can be predictive of new cases. [...] The data presented above have been shown to be consistent with a proposal for a linguistically principled account of a possible form of this disorder. This account suggests that the dysphasics do not have the normal language-learning mechanism, described by Pinker (1984), that would allow or, perhaps even compel them, to construct inflectional paradigms on the basis of regularities hypothesized from the observed linguistic evidence. The inability to construct such paradigms would result in a selective impairment of that component of grammar that encodes abstract morphology, while sparing other linguistic abilities such as the determination of argument structures.”

do primeiro: a subdivisão do próprio módulo linguístico em dois componentes relativamente autônomos – o léxico e o sistema computacional – que poderiam ser seletivamente afetados. Assim, o fato de os indivíduos disfásicos da família KE serem incapazes de conjugar adequadamente os verbos regulares (cuja flexão exige a computação de regras) e o fato de conseguirem manipular corretamente os verbos irregulares (cujo manejo depende da recuperação lexical de formas verbais armazenadas no dicionário mental) eram vistos pelas autoras como evidências de um distúrbio que corrompia o sistema computacional sem, no entanto, afetar o acesso lexical.

Em 1990, Myrna Gopnik já havia publicado na *Nature* um resumo da pesquisa que vinha desenvolvendo com a família KE. Nesta breve comunicação, intitulada “*Feature-blind grammar and dysphasia*”, Gopnik (1990) adianta os resultados de alguns testes e apresenta a tese principal do trabalho que viria a ser publicado no ano seguinte.

Como os déficits são aparentes em todos os aspectos da linguagem, suas raízes provavelmente estão na gramática subjacente, e não em um sistema de processamento periférico. Como as habilidades de linguagem que não são prejudicadas são pelo menos tão complexas quanto aquelas que são afetadas, é improvável que o déficit subjacente seja um déficit cognitivo geral. (GOPNIK, 1990: 715)⁸⁸.

As oposições não demoraram a vir. Ainda em 1990, Vargha-Khadem & Passingham publicam na *Nature* uma resposta ao texto de Gopnik (1990). Segundo os autores, que faziam parte do grupo de pesquisadores que acompanhavam a família KE no *UCL Great Ormond Street Institute of Child Health* de Londres, os indivíduos afetados apresentavam problemas em aspectos linguísticos que não estavam relacionados apenas à competência computacional. Eles, por exemplo, cometiam erros semânticos em tarefas de nomeação e tinham um desempenho abaixo do esperado em testes de vocabulário receptivo. Também foram identificadas dificuldades em tarefas de repetição de fonemas, palavras, pseudopalavras e sentenças. E mesmo dentro do sistema computacional, havia evidências de que o comprometimento gramatical era pervasivo e não se restringia ao processamento de marcadores morfológicos (de tempo, plural, pessoa etc.). Em um dos testes mencionados pelos autores, os KE afetados demonstraram um déficit na compreensão de uma série de construções que exigiam um processamento sintático mais complexo, como passivas

⁸⁸ “Because the deficits are apparent in all aspects of language, their roots probably lie in the underlying grammar rather than in a peripheral processing system. Because the language skills that are not impaired are at least as complex as those which are, it is unlikely that the underlying deficit is one of cognition as such.”

reversíveis⁸⁹, orações relativas e orações encaixadas. “*Gopnik se concentrou em apenas um aspecto do transtorno, e é impreciso concluir que o problema de fala e linguagem decorre simplesmente de uma deficiência na manipulação de marcadores de traços.*” (VARGHA-KHADEM & PASSINGHAM, 1990: 226)⁹⁰.

Em uma correspondência publicada na mesma edição da *Nature*, Paul Fletcher (1990) também se mostra cético com relação à hipótese de Gopnik (1990). Primeiro, porque, assim como Vargha-Khadem & Passingham (1990) haviam mencionado, uma série de outras pesquisas com a família KE apresentavam evidências de que o déficit apresentado pelos membros afetados não se limitava à morfologia flexional, atingindo também habilidades como complementação verbal, compreensão e uso de advérbios temporais e construção de sentenças complexas. Segundo, porque, em um dos testes de flexão citados pelo autor, apesar de o grupo disfásico apresentar um desempenho médio inferior ao do grupo controle, metade dos afetados atingiu um *score* de 60% de acertos em tarefas de manipulação de marcadores morfológicos (formação de plural e de passado regular, por exemplo). Como explicar um índice de acertos tão elevado, relativo a um número tão significativo de sujeitos afetados, se eles eram *feature-blind* (destaque do autor)? Além disso, como já mencionamos, é muito comum que portadores de DEL apresentem déficits fonológico-fonéticos responsáveis pelo apagamento de determinados itens sufixais (como o “-s” e o “-ed” do inglês).

Formas gramaticais como plural, pretérito e terceira pessoa do singular são particularmente vulneráveis, em inglês, em indivíduos com problemas de produção fonológica. A habilidade fonológica dos sujeitos foi revisada e eliminada como fator de comprometimento de alguns ou

⁸⁹ Uma “passiva reversível” é uma sentença em que o sintagma nominal que ocupa a posição de sujeito pode ser trocado pelo sintagma nominal que exerce a função (temática) de agente, sem que a sentença resultante se torne ilógica (embora o seu significado seja o oposto da sentença original). Exemplo: *A onça foi engolida pela cobra.* vs. *A cobra foi engolida pela onça.* Nas construções passivas irreversíveis, a reversão sujeito-agente produz uma construção anômala do ponto de vista lógico-semântico. Exemplo: *A petição foi assinada pelo juiz.* vs. *#O juiz foi assinado pela petição.* Restrições semânticas impedem que os sujeitos de construções passivas irreversíveis (normalmente, sintagmas nominais marcados com o traço [– animado]) exerçam a função de agente. Essas restrições lógicas, obviamente, dependem do universo que é estabelecido como referência. Em mundos imaginários (como o de “Alice no País das Maravilhas”) seria perfeitamente aceitável a ideia de uma petição assinando um juiz.

⁹⁰ “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.”

de todos eles? Até que explicações alternativas tenham sido explicitamente consideradas, deve haver algum ceticismo sobre a tese de Gopnik. (FLETCHER, 1990: 226)⁹¹.

Retomando alguns dos contrapontos apresentados na comunicação de 1990, Vargha-Khadem *et al.* publicam em 1995 os resultados de uma pesquisa realizada com vinte e um membros da família KE (treze afetados e oito não afetados), que foram avaliados em testes de produção e compreensão de linguagem, incluindo gramática, fonologia, leitura e escrita. A média de idade do grupo de afetados era de 24,4 anos (com idades entre 6 e 75), enquanto a idade média do grupo controle era de 14,1 anos (com idades entre 8 e 21). Com exceção de um único indivíduo, todos os familiares investigados eram os mesmos que haviam sido testados por Gopnik & Crago (1991).

Os autores analisaram os resultados considerando dois recortes. No primeiro, que incluía todos os indivíduos da amostra, os membros afetados obtiveram desempenhos significativamente piores do que os indivíduos do grupo controle em todos os “testes de fala e linguagem” (aqueles que não envolviam operações gramaticais), exceto em dois (nomear objetos e apontar figuras a partir de uma palavra dita pelo entrevistador). Quando a segunda geração da família foi excluída da análise contrastiva (segundo os autores, para melhor equacionar as idades dos membros afetados e não afetados), a assimetria dos resultados foi ainda mais expressiva e os membros afetados apresentaram sinais importantes de comprometimento em todos os testes, exceto com relação à nomeação de objetos. Em especial, foram atestados déficits significativos relacionados ao componente fonológico da linguagem – uma característica comum a vários subtipos de DEL – e os indivíduos afetados tiveram desempenhos insuficientes em tarefas que envolviam desde a repetição de pseudopalavras (uma competência relacionada à alça fonológica da memória de trabalho, sistema dedicado ao armazenamento e à recuperação de informações fonêmicas durante um curto período de tempo) até a deleção de seus fonemas iniciais (uma competência relacionada à capacidade de decompor entidades complexas e de discriminar segmentos sonoros com valor linguístico). Os resultados abaixo do esperado em outros testes – adição de fonemas, repetição de sentenças, decisão lexical, produção de rimas, entre outros – apontavam para a

⁹¹ “Grammatical forms like plural, past tense and third person singular are particularly vulnerable, in English, in individuals with phonological production problems. Was the phonological ability of the subjects reviewed, and eliminated as a factor in the impairment of some or all of them? Until alternative explanations have been explicitly considered, there must be some scepticism about Gopnik's thesis.”

existência de um distúrbio linguístico muito mais difuso do aquele sugerido por Gopnik & Crago (1991).

Os membros afetados também tiveram desempenhos contrastivamente piores em todos os testes de computação gramatical aplicados, com exceção de um subteste que envolvia julgamento de gramaticalidade de palavras marcadas morfológicamente (ver logo abaixo). Dos quatro testes de gramática, três avaliavam especificamente a capacidade de processar flexões: produção de formas verbais flexionadas em diferentes tempos (*tense production*); produção de marcadores morfológicos (incluindo flexão e derivação) em palavras (“*This creature is smaller than this one, but this creature must be the _____. [smallest]*”) e pseudopalavras (“*This creature is ponner than this one, but this creature must be the _____. [ponnest]*”) e julgamento de gramaticalidade de marcadores morfológicos de palavras (“Qual é a sentença correta? ‘*Planes are faster than trains, or planes are fastest than trains*’.”) e pseudopalavras (“*Planes are donker than trains, or planes are donkest than trains*.”). O outro teste (TROG: *Test for Reception of Grammar*) avaliava não apenas o conhecimento morfossintático, mas também o (re)conhecimento de outras estruturas gramaticais e de outros processos computacionais, como a inserção de sentenças sob a forma de orações relativas. Também nesse teste, os membros afetados e não afetados diferiram significativamente, indicando que a deficiência de processamento dos indivíduos afetados se estendia para além da computação de flexões e derivações, incluindo outros aspectos da competência gramatical.

O teste de produção verbal (*tense production*) desenhado pelos pesquisadores tinha o objetivo de avaliar a capacidade dos participantes mudarem as formas verbais regulares e irregulares do presente para o passado (“*Everyday I wash my clothes. Yesterday I _____ my clothes.*” [washed]) e do passado para o presente (“*Yesterday I washed my clothes; every day I _____ my clothes.*” [wash]). Os autores estavam particularmente interessados em reanalisar o déficit morfossintático seletivo que Gopnik & Crago (1991) haviam identificado nos membros afetados da família KE, que pareciam ser incapazes de gerar adequadamente formas flexionadas de verbos regulares, mas que mantinham relativamente intacta a capacidade de produzir formas apropriadas de verbos irregulares. A base para essa diferença de desempenho, como já mencionamos, foi atribuída ao tipo de operação dispensada à mudança de forma dos verbos irregulares, que não são regidos por regras de acoplagem morfológica, mas sim por diretrizes de acesso lexical.

Os resultados indicaram que a média de acertos do grupo de não afetados foi de 37,43 (em 40 possíveis) contra uma média de 19,91 entre os afetados, demonstrando que o distúrbio tem fortes implicações sobre a morfologia flexional; no entanto, o comprometimento foi igualmente evidente no processamento de formas regulares e irregulares (cf. a Fig. 5.1a). Segundo os autores, essa deficiência de formas verbais irregulares não poderia ser fruto de uma ausência de itens no léxico, já que verbos irregulares são tão frequentes quanto os regulares nos contextos normais de uso do inglês e seria improvável que os KE afetados não tivessem sido expostos às formas adequadas, ainda mais considerando-se que os indivíduos controle, que tiveram desempenhos significativamente superiores, eram da mesma família e conviviam no mesmo ambiente linguístico. Esses resultados, portanto, contrariavam a suposta preservação do sistema de processamento dos verbos irregulares. Além disso, contrariando a tese de Gopnik & Crago (1991), os pesquisadores identificaram que 41% dos erros cometidos pelos indivíduos afetados foram de supergeneralização da regra “Verbo + -ed”, levantando algumas suspeitas com relação à suposta incapacidade desses sujeitos de abstrair e adquirir regras gramaticais. Note-se, entretanto, que há uma diferença fundamental entre uma supergeneralização “instintiva,” que é fruto de uma sobregeração de formas produzidas por uma gramática mental que está mapeando o *input* linguístico de modo a encontrar nele determinadas regularidades, e uma supergeneralização “consciente”, que é fruto de estratégias cognitivas mais gerais (relacionadas à inferência, por exemplo). Assim como uma das participantes da pesquisa de Gopnik & Crago (1991) sussurrava para si mesma “*add an -s*” durante uma tarefa de formação de plural, os KE afetados poderiam estar aplicando uma regra do tipo “*add an -ed*” não como resultado de um mecanismo gramatical internalizado de formação do passado regular, mas sim como uma instrução memorizada de forma consciente.

O déficit articulatório – um dos traços mais evidentes dos membros afetados da família KE e que não havia recebido atenção no estudo de Gopnik & Crago (1991) – também foi avaliado por Vargha-Khadem *et al.* (1995). Para examinar até que ponto a fala comprometida dos KE poderia estar sendo causada por um distúrbio motor subjacente, responsável por uma desarticulação de movimentos orais e faciais de todos os tipos, foram aplicadas duas modalidades de testes de praxia orofacial. No primeiro, os participantes precisavam executar comandos verbais dados pelo examinador (por exemplo, “*imite o animal X*”; “*morda o lábio*”; “*feche o olho esquerdo*”; “*estique a língua, lamba o lábio superior e estale os lábios*”). Os resultados revelaram diferenças significativas entre o grupo de afetados e o grupo controle. Para evitar que uma possível dificuldade em compreender os comandos

estivesse interferindo na execução correta da tarefa, foram aplicados testes de imitação, nos quais os participantes precisavam imitar os movimentos faciais do examinador. Foram avaliados três grupos de movimentos orofaciais: movimentos simples (como abrir a boca); movimentos complexos (projetar o lábio inferior e a mandíbula, por exemplo); movimentos paralelos (como abrir a boca vocalizando e projetando a língua ao mesmo tempo) e movimentos sequenciais (como cerrar os lábios → abrir a boca → mostrar a língua). Os resultados mostraram novamente uma disparidade de desempenho entre os indivíduos não afetados e os afetados, que conseguiam realizar movimentos simples com um pouco mais de desenvoltura, mas que apresentavam dificuldades cada vez mais salientes à medida que o grau de complexidade do movimento aumentava (cf. a Fig. 5.1b). Esses resultados indicavam que o déficit práxico verbal (relacionado à desarticulação dos movimentos da fala) era na realidade um reflexo de um déficit práxico orofacial subjacente.

Contrapondo-se em parte à caracterização fenotípica de natureza exclusivamente linguística proposta por Gopnik & Crago (1991) e ampliando a hipótese originalmente apresentada por Hurst *et al.* (1990), Vargha-Khadem *et al.* (1995) adotam uma posição que tenta abarcar todo o quadro de déficits do distúrbio identificado na família KE.

Nossas investigações da família KE indicam que o distúrbio hereditário tem um fenótipo amplo que transcende a geração prejudicada de regras morfossintáticas e inclui um notável comprometimento articulatorio, bem como defeitos nas funções práxicas, intelectuais, linguísticas e orofaciais em geral. Embora a evidência dessa família não forneça suporte para a existência de “genes gramaticais”, suas dificuldades linguísticas constituem uma parte proeminente de seu fenótipo. Investigações dos correlatos neurais e genéticos de seu distúrbio poderiam, portanto, fornecer pistas importantes para algumas das bases das faculdades humanas primárias de fala e linguagem. (VARGHA-KHADEM *et al.*, 1995: 930)⁹².

Esse “fenótipo amplo”, apesar da vantagem de abarcar quase todo o espectro de traços e sintomas do distúrbio da família KE, tem a desvantagem de parecer pouco preciso, uma vez que não delimita o déficit nuclear do transtorno (aquele que poderia estar eventualmente na base de todos ao outros). A fim de reduzir esse fenótipo ao seu defeito central, Vargha-

⁹² “Our investigations of the same KE family indicate that the inherited disorder has a broad phenotype which transcends impaired generation of syntactical rules and includes a striking articulatory impairment as well as defects in intellectual, linguistic, and orofacial praxic functions generally. Although the evidence from this family thus provides no support for the existence of “grammar genes,” their linguistic difficulties do constitute a prominent part of their phenotype. Investigations of the neural and genetic correlates of their disorder could therefore uncover important clues to some of the bases of the primary human faculties of speech and language.”

Khadem *et al.* publicam um novo trabalho em 1998, apresentando os resultados de uma série de testes aplicados à família KE, em conjunto com os padrões neurológicos observados durante a realização dessas tarefas.

O estudo se baseou em duas tarefas de coordenação oromotora: (i) repetição de palavras e pseudopalavras; (ii) execução de movimentos orofaciais (simultâneos e sequenciais). Em ambas as tarefas, houve uma assimetria expressiva entre o desempenho dos indivíduos controle e o dos afetados (Fig. 5.12). Durante a resolução da tarefa (i) – que envolvia o processamento articulatório da fala, uma atividade complexa que exige um alto grau de precisão muscular – a atividade cerebral dos participantes foi monitorada por PET e por imagens de ressonância magnética⁹³. Em indivíduos afetados, foram identificadas anormalidades *funcionais* em áreas motoras corticais e subcorticais do lobo frontal e também anormalidades *estruturais* em várias dessas mesmas áreas, particularmente no núcleo caudado, que se apresentava anormalmente pequeno em ambos os hemisférios. Para Vargha-Khadem *et al.* (1998), a observação na tarefa (i) de um padrão incomum de atividade neurológica em áreas corresponsáveis pelo processamento motor, associada aos resultados dos testes de execução de movimentos orofaciais na tarefa (ii), indicavam que a dispraxia verbal dos KE era derivada de um déficit de praxia orofacial e de articulação sequencial causado por problemas neurológicos. Dessa forma, o traço fenotípico nuclear do distúrbio passou a ser caracterizado como um déficit motor orofacial. Uma mutação genética na região

⁹³ A tomografia por emissão de pósitrons (PET, do inglês, *Positron Emission Tomography*) e a ressonância magnética funcional (fMRI, do inglês, *Functional Magnetic Resonance Imaging*) são duas técnicas não invasivas que permitem observar detalhadamente o funcionamento do cérebro humano durante a realização de tarefas cognitivas. Ferramentas como essas permitem mapear os centros neuronais que têm sua atividade metabólica aumentada durante o processamento linguístico: “A maioria dessas técnicas (em particular, a PET e a fMRI) se baseiam no pressuposto de que um incremento da atividade neuronal de uma determinada região cerebral, como consequência de sua participação em alguma tarefa relacionada com o processamento de informação linguisticamente relevante, implica um aumento paralelo do fluxo sanguíneo para essa área, com o objetivo de garantir um suprimento adequado de oxigênio e nutrientes que seja capaz de sustentar o incremento da atividade metabólica associada a tal tarefa [...]. A conclusão fundamental obtida a partir dos estudos realizados mediante estas técnicas é substancialmente a mesma, a saber, que os centros encarregados pelo processamento da linguagem não são homogêneos, mas se organizam como uma rede difusa de subcentros responsáveis pelo processamento de componentes específicos da linguagem. Em termos gerais, esses subcentros neuronais podem ser descritos de forma mais adequada em termos de sua especialização linguística, de modo que os diferentes componentes da linguagem (fonologia, semântica, sintaxe) tenham aproximadamente um correlato estrutural, mas acima de tudo funcional, no que diz respeito à organização do córtex cerebral.” (BURRACO, 2012: 7-8).

cromossômica SPCH1 (que havia sido identificada há pouco tempo por Fisher *et al.*, 1998) havia condicionado o desenvolvimento anormal de várias áreas do cérebro que pareciam ser críticas tanto para os movimentos orofaciais quanto para a articulação sequencial (levando, conseqüentemente, a uma degeneração acentuada da fala). Essa tese – segundo a qual o principal traço fenotípico dos indivíduos afetados da família KE é o comprometimento de sua motricidade orofacial – seria retomada em Vargha-Khadem *et al.* (2005), agora com o gene *FOXP2* devidamente identificado, mapeado e sequenciado.

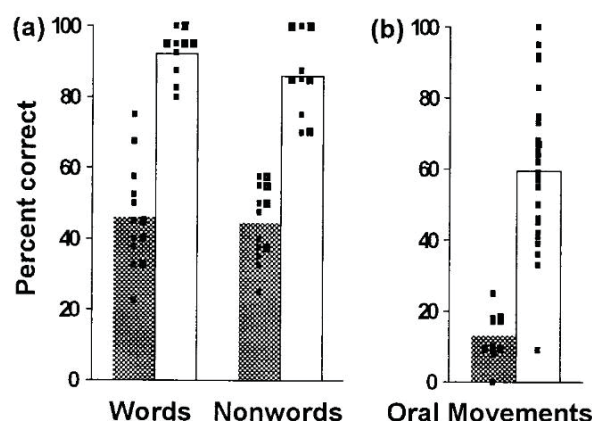


FIGURA 5.12. (a) Repetição de palavras e pseudopalavras. As barras indicam a porcentagem média de respostas corretas para o grupo de familiares afetados (barra preenchida; $n = 13$) e para o grupo de não afetados (barra vazada; $n = 10$). (b) Movimentos orofaciais simultâneos e sequenciais realizados sob comando. As barras indicam a porcentagem média correta para o grupo de familiares afetados (barra preenchida; $n = 11$) e para o grupo controle normal (barra vazada; $n = 52$). Em (a) e (b), os quadradinhos representam os desempenhos individuais dos participantes. Imagem extraída de VARGHA-KHADEM *et al.* (1998).

Apesar de os déficits motores terem sido elencados como centrais na caracterização do distúrbio da família KE, os déficits morfossintáticos identificados inicialmente por Gopnik e colaboradores continuaram a ser investigados. Em 2002, uma nova bateria de testes foi aplicada à família KE por Kate Watkins e colaboradores. A pesquisa contou um grupo de treze indivíduos afetados da família KE e dois grupos controle: membros familiares não afetados e pacientes com afasia adquirida (resultante de acidente vascular cerebral). Aspectos da morfologia flexional e derivacional foram novamente avaliados em um dos testes. Os participantes precisavam completar a segunda de duas sentenças relacionadas, de modo que a primeira continha a palavra que deveria ser modificada morfologicamente para completar a segunda. Para facilitar a compreensão, o examinador mostrava ao participante uma figura que ilustrava a situação narrada. Foram utilizadas estruturas com palavras (“*Look at how small*

these elephants are. This one here must be the _____?” [smallest]) e pseudopalavras (“*This creature is snozzing. We call him a _____?*” [snozzler]). A resolução dos problemas exigia a manipulação de morfemas e regras derivacionais e flexionais (verbais e nominais). Confirmando o que havia sido observado em outros estudos, os resultados indicaram uma diferença expressiva de desempenho entre o grupo de afetados e o grupo de não afetados; entre os afetados e os afásicos, por sua vez, não houve diferença do ponto de vista estatístico (Fig. 5.13).

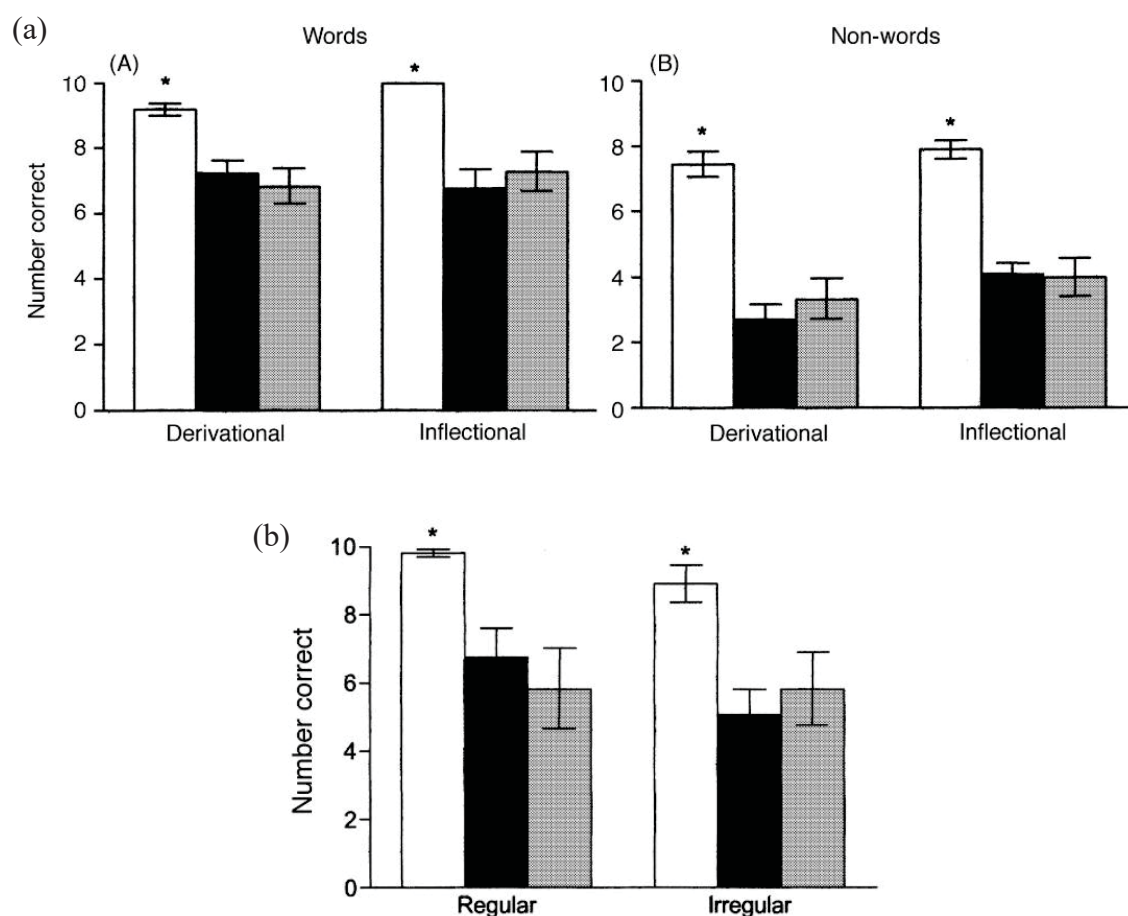


FIGURA 5.13. (a) Produção de morfologia derivacional e flexional para palavras (A) e pseudopalavras (B). (b) Flexão de verbos regulares e irregulares (produção do passado). As barras representam a pontuação média dos não afetados (barra vazada), afetados (barra escura) e afásicos (barra cinza). Os segmentos do topo indicam a margem de erro, para mais e para menos. O asterisco indica diferença significativa entre o grupo não afetado e os grupos afetado e afásico. Extraído de WATKINS *et al.* (2002).

A partir desse conjunto acumulado de evidências, fica claro que os indivíduos que apresentam uma versão defeituosa do *FOXP2* apresentam problemas de natureza morfossintática relacionados tanto à flexão quanto à derivação. O mecanismo exato por trás desse fenômeno permanece desconhecido, mas uma das hipóteses é de que a ausência de

doses adequadas da proteína FOXP2 em determinadas fases da embriogênese impede o desenvolvimento normal de áreas do cérebro ligadas à computação linguística (como a área de Broca). No entanto, a construção adequada de cada agrupamento neuronal corresponsável por algum tipo de função linguística depende de inúmeros genes e seria uma ilusão imaginar que o *FOXP2* poderia ser o fator único responsável pela competência morfossintática. Indivíduos portadores de afasia de Broca e de outras formas de DEL, por exemplo, exibem déficits morfossintáticos semelhantes aos dos KE (cf. Rice & Wexler, 1996), mesmo possuindo na maioria das vezes duas cópias intactas do *FOXP2*.

As evidências apresentadas pelos estudos que selecionamos neste trabalho (em conjunto com os resultados de outras pesquisas) mostram que o fenótipo da família KE condicionado pela mutação do *FOXP2* não se restringe a um “transtorno sensório-motor” (SHRIBERG et al., 2006) ou a um déficit morfossintático. Como vimos em alguns dos testes mencionados, outros componentes da linguagem parecem estar igualmente corrompidos. A taxa de erro que os indivíduos afetados exibem ao manipular formas irregulares (que são armazenadas diretamente no léxico mental), por exemplo, é semelhante à observada no caso de formas regulares (que são o resultado da aplicação de regras combinatórias), indicando que há problemas tanto na lexicalização quanto na capacidade de recuperar elementos do léxico. Em termos qualitativos, os indivíduos afetados obtiveram resultados abaixo do esperado em testes que avaliaram a capacidade de nomear objetos a partir de desenhos; em termos quantitativos, o tempo de latência de resposta a estes testes também foi maior (WATKINS et al. 2002). Em um dos testes aplicados por Watkins et al. (2002), a fluência lexical do grupo afetado (i.e., a capacidade de gerar o maior número possível de itens em um determinado período de tempo), apesar de ter sido superior à do grupo de afásicos, foi significativamente inferior à do grupo não afetado, tanto em condições de restrição semântica (“diga o nome de frutas”) quanto fonêmica (“diga palavras que comecem com a ‘letra m’”). Essa menor fluência também foi observada no teste de escrita fonêmica (“escreva palavras que comecem com a ‘letra s’”), o que indicaria que o distúrbio existente nesses indivíduos não é de natureza exclusivamente articulatória (mesmo assim, é importante ressaltar que a capacidade de articulação linguística parece ser um ponto de corte fundamental na delimitação do fenótipo. Nesse mesmo estudo, Watkins et al. (2002) demonstram que o teste mais confiável para realizar a distinção entre indivíduos afetados e não afetados na família KE é o teste de repetição de pseudopalavras com padrões articulatórios complexos, mais especificamente, encontros consonantais).

O espectro amplo de déficits linguísticos dos indivíduos afetados da família KE (que apresentam desde comprometimentos na memória de trabalho fonológica até dificuldades em processar semanticamente sentenças com orações encaixadas) e a transversalidade de seus problemas (fonêmicos, morfológicos, sintáticos, semânticos, expressivos e discursivos), somados ao seu déficit articulatório oromotor, desde o início representaram um desafio para a caracterização precisa de seu fenótipo.

A identificação do déficit nuclear do distúrbio – do qual todos os outros déficits poderiam derivar ou ser tributários – se transformou em uma questão tão importante quanto a própria identificação do gene responsável pelo transtorno: ao mesmo tempo em que os geneticistas esquadrihavam o genoma da família KE em busca do fator molecular determinante, linguistas e neurocientistas tentavam reduzir o fenótipo à sua essência, assumindo muitas vezes hipóteses de trabalho conflitantes. A presença de um fator genético detectável na etiologia do déficit de linguagem dos KE – que, até então, faziam parte do único caso relatado de um distúrbio no desenvolvimento da linguagem causado pela mutação em um único gene – colocou essa família no centro de uma disputa entre correntes teóricas que reatualizavam o debate sobre o inatismo linguístico – afinal, pela primeira vez na história, havia a possibilidade de se comprovar materialmente a existência de uma diretriz gramatical codificada diretamente no genoma humano.

As diferentes posições assumidas pelos grupos de Gopnik e de Vargha-Khadem refletiam bem essa disputa entre correntes teóricas quanto à natureza da linguagem. Para Gopnik e colaboradores, o déficit nuclear do distúrbio era de natureza linguística e o gene mutado estaria afetando neurônios subjacentes ao programa gramatical inato da espécie humana; para o grupo de Vargha-Khadem, o déficit nuclear do distúrbio era de natureza motora e a mutação condicionante estaria afetando agrupamentos de neurônios responsáveis pela programação e execução de movimentos orofaciais. As duas correntes admitiam a coocorrência de ambos os déficits – linguístico e motriz –, mas diferiam quanto à ideia de que algum princípio gramatical inato poderia estar sendo afetado pela mutação identificada na família KE.

Hoje, para fins de caracterização fenotípica, o déficit prático oromotor, pelo próprio caráter visível de sua manifestação física, é considerado o aspecto nuclear do distúrbio apresentado pelos KE. Mas essa opção não deve ser entendida como uma negação dos déficits linguísticos que fazem parte da dimensão cognitiva do transtorno. A circunscrição do fenótipo

ao seu aspecto motor serve muito mais para caracterizá-lo de uma forma segura e menos difusa. Todas as evidências indicam que o *FOXP2*, por codificar um fator de transcrição envolvido em inúmeros processos celulares, possui uma natureza pleiotrópica, influenciando inúmeras características, estruturais e funcionais. Ao participar da construção e do funcionamento de determinados agrupamentos de neurônios do cérebro humano, o *FOXP2* se encontra etiologicamente ligado não só ao substrato material que é por ele determinado, mas à atividade abstrata que emerge dessa base eletrofisiológica, fazendo com que seu fenótipo se desdobre em uma dimensão virtual (cognitiva). A escolha de uma característica física para definir o núcleo fenotípico do distúrbio apresentado pelos KE – e, conseqüentemente, o núcleo fenotípico da mutação do *FOXP2* – decorre simplesmente do fato de que, metodologicamente, é mais simples circunscrever o concreto do que circunscrever o abstrato. Pelo espectro de déficits puramente linguísticos que emergem de suas mutações, é provável que o *FOXP2* faça parte do grupo de cofatores genéticos responsáveis pela competência linguística humana – e suspeita-se que ele seja um dos mais importantes.

Neste capítulo, optamos por um exame mais minucioso do *FOXP2* a partir do distúrbio linguístico-motor apresentado pelos KE, uma vez que essa família representa o caso paradigmático. No entanto, posteriormente foram identificados outros indivíduos e grupos familiares que apresentavam déficits semelhantes aos dos KE afetados e cuja análise genômica revelou mutações no *FOXP2*. Merecem destaque o menino CS (LAI *et al.*, 2001), fundamental na localização do *FOXP2*, e a microfamília TB (SHRIBERG *et al.*, 2006), cujos indivíduos afetados – mãe e filha – eram portadores de uma translocação cromossômica que havia interrompido a sequência codificadora do *FOXP2*, causando problemas articulatorios e linguísticos semelhantes aos observados entre os KE.

RESUMO DO CAPÍTULO E CONSIDERAÇÕES FINAIS

A história da descoberta do gene *FOXP2* está indissociavelmente ligada à história da família KE e do distúrbio de fala e linguagem que afetava quase metade de seus membros. Foi a partir da detecção do padrão de herança desse distúrbio – que parecia estar sendo causado por uma mutação dominante em um único gene – que foram iniciadas as buscas pelo fator genético condicionante, culminando com a identificação do *FOXP2* em 2001.

Localizado no cromossomo 7, o *FOXP2* pertence a um grupo de genes que regulam a ação de outros genes através da produção de fatores de transcrição – proteínas que bloqueiam ou ativam a transcrição (e, consequentemente, a expressão) de outros genes. Como todas as células de um organismo possuem essencialmente o mesmo conjunto de genes, é preciso que um mecanismo controle quais devem ser expressos ou não. Genes como o *FOXP2* permitem uma taxa diferencial de expressão gênica em cada célula, de acordo com as exigências proteicas de seu metabolismo, sua função e sua estrutura. Taxas diferenciais de expressão gênica também são as responsáveis pela diferenciação celular durante o desenvolvimento embrionário. A proteína FOXP2 está presente em altas concentrações em determinadas áreas do cérebro humano durante seu desenvolvimento e é uma das responsáveis pela diferenciação e formação de agrupamentos neuronais que, entre outras funções, estão envolvidos na programação e na coordenação dos movimentos sequenciais necessários à produção da fala. Além de estarem associados à atividade motora orofacial, esses agrupamentos de neurônios também desempenham tarefas cognitivas e podem ter uma relação direta com a regulação de determinadas propriedades da competência linguística. Assim, mutações no *FOXP2* podem desencadear um distúrbio caracterizado principalmente por (i) déficits de motricidade que afetam a articulação da fala e (ii) problemas de processamento linguístico que derivam de déficits do sistema gramatical.

No caso específico da família KE, todos os indivíduos afetados apresentavam uma cópia modificada do *FOXP2*, que havia sofrido uma mutação pontual em sua sequência de bases, resultando na síntese de uma proteína disfuncional (a outra cópia do gene permanecia intacta e era responsável pela codificação de uma proteína biologicamente ativa). Assim como em outros quadros heterozigóticos, a presença de apenas uma cópia funcional do *FOXP2* faz com que a quantidade disponível de proteínas ativas seja reduzida pela metade, o que, aparentemente, não é o suficiente para alcançar os níveis críticos necessários para a formação e o funcionamento adequados dos agrupamentos de neurônios que dependem de uma forma

ou de outra dessas proteínas. Esse déficit quantitativo, no entanto, não afeta o funcionamento de outros órgãos em que o *FOXP2* é expresso – como os pulmões e o intestino –, que conseguem exercer suas atividades apenas com a metade da dose proteica sintetizada pela cópia ativa do gene.

Além de ser expresso em outros órgãos (e não apenas no cérebro), o *FOXP2* faz parte do genoma de outras espécies (e não apenas dos seres humanos). Versões levemente diferentes desse gene podem ser encontradas em vertebrados como répteis, camundongos, aves e grandes primatas, onde exercem papéis importantes na regulação da atividade motora, entre outras funções ainda não muito bem conhecidas. As análises dos padrões de expressão cerebral do *Foxp2* em aves canoras revelam que esse gene possui um papel importante no processo de aprendizagem vocal, estando aparentemente relacionado com a capacidade inata que essas aves possuem de adquirir e produzir cantos com padrões sonoros complexos. Dada a raridade do distúrbio causado por mutações do *FOXP2* em humanos, os estudos com modelos animais – especialmente com aves canoras – parecem ser um caminho promissor para a investigação dos papéis biológicos exercidos por esse gene, particularmente com relação a propriedades que podem ser transpostas para o estudo da linguagem.

O espectro de déficits linguísticos da família KE é amplo – indo desde problemas relacionados à morfologia flexional e derivacional até dificuldades significativas em julgamentos de gramaticalidade de sentenças do inglês – e desde o início motivou discussões acaloradas acerca da real natureza do distúrbio. No entanto, apesar das dificuldades relativas à delimitação precisa do fenótipo, entre as principais características associadas à mutação do *FOXP2*, podemos elencar os seguintes sintomas (as informações entre colchetes e em itálico presentes na citação a seguir são acréscimos de nossa autoria):

Dificuldades receptivas e expressivas de vários tipos, incluindo (i) um menor nível de conhecimento lexical; (ii) menor capacidade de compreensão morfológica e morfossintática; (iii) problemas de geração morfológica, tanto de substantivos como de verbos, que afetam igualmente as formas regulares e irregulares, e que muitas vezes envolvem processos de supergeneralização (verbal); (iv) um déficit na capacidade de armazenamento de informações fonologicamente relevantes pela memória de trabalho verbal ou talvez no processo de articulação sequencial de unidades com valor fonológico; (v) uma menor capacidade de recuperação de elementos do léxico (em termos quantitativos e qualitativos); [(vi) *dificuldades em julgamentos de gramaticalidade*; (vii) *um comprometimento no processamento de estruturas sintáticas complexas (com orações encaixadas, por exemplo)*; (viii) *limitações no uso de pronomes e elementos reflexivos*; (ix) *produção de sentenças com poucos elementos*

lexicais/morfológicos (baixa MLU);] [...] e (x) uma dispraxia verbal que dá origem a um discurso falado distorcido, marcado principalmente por (a) vários problemas para a correta articulação de certos tipos de vogais e consoantes; (b) vários tipos de distorções e substituições fonêmicas; (c) [aféreses (supressão de fonemas iniciais), apócopies (supressão de fonemas finais)], epênteses [adição de fonemas], metáteses [troca interna de fonemas] e assimilações/reduções atípicas; (d) capacidade de fraseado diminuída; (e) lentidão significativa da fala; (f) um padrão de acento de sentença desviante; (g) presença de bloqueios articulatórios antes de certas consoantes oclusivas; (h) um prolongamento anormal de certos sons sonantes; e (i) uma alteração de certos parâmetros do discurso falado, incluindo sua qualidade laríngea e sua ressonância (BURRACO, 2012: 249).

Todo esse quadro de déficits fonêmicos característicos da dispraxia verbal é tributário de uma dispraxia orofacial subjacente que compromete a atividade motora da musculatura do rosto, especialmente com relação à execução de movimentos sequenciais complexos.

CAPÍTULO 6 – BREVE CONSIDERAÇÃO FINAL

No início deste trabalho, apresentamos o problema que orientou a nossa investigação.

(1) Quais são e de que modo atuam os genes (ou fatores genéticos) que determinam a faculdade da linguagem (FL)?

Mais do que fornecer uma solução para esse problema, procuramos analisá-lo a partir de um conjunto de informações e dados que julgamos importantes para uma discussão *introdutória* sobre o tema, buscando, sempre que possível, estabelecer uma ponte entre a Linguística e a Genética.

Há um pressuposto fundamental por trás do enunciado (1): a ideia de que a faculdade da linguagem (i.e., o programa subjacente que nos permite adquirir e usar uma língua) é determinada geneticamente. Esta pode parecer uma afirmação forte, mas a hipótese alternativa, segundo a qual a faculdade da linguagem *não é* determinada geneticamente, por tudo que discutimos até aqui, parece-nos tão ou mais forte ainda. Mesmo em modelos conexionistas – que se fundamentam na ideia de que a língua (enquanto realidade mental) é o *output* de uma computação realizada por unidades neurocognitivas de processamento da experiência linguística que *não* são orientadas por um programa gramatical inato –, a predisposição genética para a aquisição da linguagem está de alguma forma presente, mesmo que indiretamente, já que esse aparato “multifuncional” de aquisição possui necessariamente uma contraparte no DNA. A diferença fundamental é que em tais modelos não se admite a existência de princípios inatos exclusivamente linguísticos – ou mesmo a existência de neurônios especializados em processar informação linguística (hipótese inatista fraca) – mas apenas a existência de princípios inatos gerais de processamento cognitivo (de toda forma, também condicionados geneticamente). Discutimos alguns dos problemas desse tipo de abordagem no capítulo 3.

Neste trabalho, portanto, assumimos que há um programa gramatical inato – uma GU – que faz parte da “configuração *default*” da espécie humana. Ao longo do empreendimento gerativista, a conceitualização da GU foi sofrendo uma série de formatações, sempre na tentativa de encontrar os seus princípios mínimos. Acompanhar as diversas fases da GU

enquanto modelo teórico é também acompanhar a história de uma tentativa de eliminar de seu corpo de diretrizes tudo aquilo que, mesmo sendo importante, é acessório, ficando apenas com o essencial. Essa redução teórica da GU ao seus elementos mínimos não corresponde apenas ao desejo da GGT de tornar o modelo mais “elegante” – mais simples e igualmente preditivo – mas é motivada pela própria natureza da GU enquanto programa geneticamente determinado. Do ponto de vista biológico e evolutivo, é extremamente complicado explicar como princípios altamente especializados – como os Princípios A, B e C da Teoria de Regência e Ligação (TRL) – podem estar codificados no DNA da espécie humana. Um modelo detalhado como o da TRL, com os seus múltiplos módulos e princípios, explica adequadamente uma série de fenômenos gramaticais, mas parece excessivamente “pesado” enquanto modelo de representação mental. A suspeita é de que a “GU real” seja constituída por princípios e operações subjacentes mais simples, estes sim codificados de alguma forma no genoma de nossa espécie. Um dos objetivos do Programa Minimalista (PM) é exatamente construir um modelo formal mais ajustado de gramática universal, com menos peças e premissas e com maior alcance e poder explicativo.

O Programa Minimalista assume como condição limite da investigação que as operações básicas da GU são simples e que as complexidades observadas nas línguas naturais são o resultado de interações entre subsistemas simples. [...] Dado o surgimento tardio da Faculdade da Linguagem (FL) em humanos, é evolutivamente natural que a FL importe operações de outras partes do sistema cognitivo. Isso sugere mais uma traço de “naturalidade”, a saber, generalidade. Operações e princípios em ação em outras partes da economia cognitiva são recursos naturais para computações linguísticas. Outra marca de “naturalidade” é a atomicidade das operações computacionais. Merge (juntar duas expressões) e Copy (duplicar uma expressão) são razoavelmente considerados como operações computacionalmente “atômicas”. Eles contrastam com regras específicas de linguagem mais complexas (como a de construção de passivas) que são analisadas como combinações de operações mais básicas. Essa concepção de “simples” e “atômica” acena para a implementação de operações semelhantes no cérebro. Quaisquer operações que os gramáticos proponham devem, em última análise, ser incorporadas aos circuitos cerebrais. É razoavelmente claro como alguém pode construir um circuito de mesclagem ou cópia, e essa é uma das razões pelas quais operações primitivas como essas são atraentes. [...]. Em suma, a FL será natural se for baseada em princípios e operações que promovam a eficiência computacional, que sejam construídas a partir de partes cognitivamente genéricas e atômicas, e que sejam básicas o suficiente para serem (plausivelmente) incorporadas em circuitos neurais. Evidentemente, mesmo com essas considerações, resta muito espaço para visões divergentes sobre como interpretar essas diretrizes e, não surpreendentemente, há um grande conjunto de potenciais candidatos para o inventário de operações e princípios básicos. No entanto, acredito que essas diretrizes podem desempenhar

um papel mais do que retórico na construção e avaliação de propostas gramaticais. Mais concretamente, acredito que a busca por operações e princípios simples sugere um projeto minimalista interessante: a construção de modelos gramaticais baseados em um pequeno inventário de operações e princípios que são ao mesmo tempo evolutiva e neurologicamente plausíveis e dos quais as propriedades básicas das gramáticas das línguas naturais podem ser derivadas qualitativamente (HORNSTEIN, 2009: 1; 3-4)⁹⁴.

Hornstein (2009) está interessado em apresentar algumas diretrizes do PM para a construção de modelos de gramática universal (FL) que sejam baseados em princípios e operações *biologicamente* possíveis, tanto em termos neurológicos quanto evolutivos. Assim, um dos objetivos do PM – talvez o mais desafiador – é propor uma teoria gramatical que integre as propriedades da FL ao substrato neurofisiológico que a realiza no cérebro humano (Problema de Broca) e à história evolutiva do *Homo sapiens* (Problema de Darwin-Wallace). Como comentamos logo no início de nosso trabalho, parece-nos indispensável que a teoria gramatical também integre as propriedades da FL ao programa genético de nossa espécie. Nesse sentido, a redução da FL a operações e princípios mínimos e naturais parece promissora

⁹⁴ “[...] The Minimalist Program [...] takes it as a boundary condition on inquiry that the basic operations of UG be simple and that the attested complexities of natural language be the result of the interactions of simple subsystems. [...] Given the rather late emergence of FL in humans it is evolutionarily natural that FL should import operations from other parts of the cognitive system. This suggests one more mark of “naturalness,” namely generality; operations and principles at work in other parts of the cognitive economy are natural resources for linguistic computations. A further mark of the “natural” is the “atomicity” of the computational operations. Merge (join two expressions) and copy (duplicate an expression) are reasonably taken as computationally “atomic” operations. They contrast with more complex language specific rules like “passive” which are reasonably analyzed as compiled combinations of more basic operations. This conception of “simple” and “atomic” casts a furtive glance towards implementation in brain like material. Whatever operations grammarians propose must ultimately be embedded in brain circuitry. It is reasonably clear how one could build a merge or copy circuit, and this is one reason that primitive operations like these are attractive. [...] In sum, FL will be natural if it is based on principles and operations that promote computational tractability, that are built from parts that are cognitively general and atomic, and that are basic enough to be (plausibly) embodied in neural circuitry. As should be evident, even given the desiderata above, there remains plenty of room for diverging views on how to interpret these guidelines and, not surprisingly, there is a large pool of potential candidates for the inventory of basic operations and principles. Nonetheless, I believe that these guidelines can play a more than rhetorical role in the construction and evaluation of grammatical proposals. More concretely, I believe that the search for simple operations and principles suggests an interesting minimalist project: the construction of grammatical models based on a small inventory of operations and principles that are at once evolutionary and neurologically plausible and from which the basic properties of natural language grammars can be qualitatively derived.”

do ponto de vista da investigação genômica, uma vez que esses processos precisam ser incorporados aos circuitos cerebrais, criando um “rastros” para o mapeamento dos genes responsáveis.

A fim de estabelecer os seus componentes e as suas operações nucleares, a FL foi conceitualizada de duas formas. A *faculdade da linguagem em sentido amplo* (FLB: *Faculty of language – broad sense*) é formada pela FLN (o núcleo computacional do sistema) e por outros componentes, incluindo o sistema sensorio-motor e o sistema conceitual-intencional que, em conjunto, dotam a nossa espécie do poder único de expressar uma gama ilimitada de pensamentos através de sequências estruturadas de sons ou sinais visuais. Alguns desses componentes (memória, sistema perceptivo, sistema vocalizador etc.) são amplamente compartilhados com outras espécies. Como vimos no primeiro capítulo, as abelhas não têm apenas a capacidade de calcular a distância e a direção da fonte de alimento, mas também a memória necessária para recuperar essas informações no momento em que transmitem a mensagem para as companheiras. Outros componentes da FLB são raros ou talvez únicos, e nos distinguem dos outros animais e de nossos ancestrais hominídeos. Isso inclui nossa capacidade de articular expressões vocais complexas e nosso desejo de compartilhar informações livremente com os outros membros do grupo (desconhecemos animal mais ~~fofoqueiro~~ comunicativo do que o *homo sapiens*). Aparentemente, somos os únicos primatas dotados de um analisador sintático inato capaz de extrair hierarquias complexas de sequências acústicas ou visuais através de operações simples e ultravelozes – mas compartilhamos versões desse *parser* com aves construtoras de melodias (com as quais partilhamos também versões do *FOXP2*, que aparentemente exerce um papel importante no aprendizado vocal dessas aves). Todos esses componentes da FLB dados como exemplo possuem, obviamente, correlatos neurológicos e são geneticamente determinados. A FL só emerge da integração ajustada de todos eles. No entanto, a capacidade de computar estruturas, pela própria natureza de sua operação, constitui o núcleo da FL.

A *faculdade da linguagem em sentido estrito* (FLN: *Faculty of language – narrow sense*) é o sistema computacional responsável pela geração de representações que são enviadas para o sistema sensorio-motor e para o sistema conceitual-intencional. Operações como Merge e propriedades como a recursividade fazem parte desse núcleo gramatical, responsável pela infinitude discreta da FL. Tais operações e princípios são suficientemente simples e primitivos para serem mapeados em áreas do cérebro implicadas em computações linguísticas.

O inventário de operações e princípios que constitui a FLN pode variar minimamente de um modelo gramatical para outro, mas deve conter apenas as ferramentas necessárias para a geração otimizada de estruturas que serão lidas pelos sistemas de desempenho. Pelo aparecimento relativamente recente da FL no curso evolutivo da espécie humana, supõe-se que a FLN tenha cooptado operações de outros sistemas cognitivos e que sua natureza única tenha emergido de uma combinação de operações e princípios transversais associados a um número pequeno de inovações.

Como vimos na discussão sobre o papagaio Alex, que era capaz de realizar uma operação combinatória entre dois itens lexicais (*blue key*), mas que não era capaz de inserir a expressão gerada em uma nova combinação, a recursão é uma das propriedades básicas da FLN, sem a qual as operações mínimas do sistema não conseguiriam gerar estruturas mais e mais complexas. Apesar de ser essencial à computação linguística, a recursão também é utilizada para resolver outros problemas computacionais, como quantificação de números. Objetos mentais virtualmente infinitos como o conjunto dos números naturais são diretamente derivados desse moto-contínuo que permite a repetição de uma mesma operação básica (nesse caso, a adição de uma unidade). Em termos genéticos e evolutivos, a questão que se coloca é saber se a recursão surgiu como uma inovação no cérebro humano, sendo causada por uma mutação e aparecendo como a peça que faltava para que a FLN girasse e a FL emergisse ou se ela seria um princípio computacional já preexistente – eventualmente presente em outros animais – que se integrou à FLN assim que esse sistema alcançou evolutivamente as condições suficientes para utilizá-la.

Se encontrarmos evidências de recursão em animais, mas em um domínio não comunicativo, é mais provável que identifiquemos os mecanismos subjacentes a essa capacidade e as pressões seletivas que levaram a ela. Essa descoberta, por sua vez, abriria a porta para outro conjunto de quebra-cabeças: por que os humanos, mas nenhum outro animal, usaram o poder da recursão para criar um sistema de comunicação aberto e ilimitado? Por que nosso sistema de recursão opera em uma gama mais ampla de elementos ou entradas (por exemplo, números, palavras) do que outros animais? Uma possibilidade, consistente com o pensamento atual nas ciências cognitivas, é que a recursão em animais representa um sistema modular projetado para uma função específica (por exemplo, navegação) e impenetrável em relação a outros sistemas. Durante a evolução, o sistema de recursão modular e altamente específico de domínio pode ter se tornado penetrável e de domínio geral. Isso abriu caminho para os humanos, talvez de forma única, aplicarem o poder da recursão a outros problemas. Essa mudança de domínio específico para domínio geral pode ter sido guiada por pressões seletivas particulares, exclusivas de nosso passado evolutivo, ou como consequência (subproduto) de outros tipos de reorganização

neural. De qualquer forma, essas são *hipóteses testáveis* [grifo nosso], um refrão teórico que destaca a importância das abordagens comparativas da faculdade da linguagem. (HAUSER *et al.*, 2002: 1578⁹⁵).

Se uma hipótese não pode ser testada quanto à sua falsidade, ela não pode ser considerada minimamente científica (apesar de, paradoxalmente, poder estar certa). O texto de Hauser *et al.* (2002), que pertence à corrente biolinguística que surgiu no seio do Programa Minimalista, exemplifica bem a preocupação do PM em propor hipóteses de trabalho testáveis sobre a natureza da faculdade da linguagem – hipóteses que podem (e devem) ser confrontadas com os mecanismos biológicos subjacentes. Assim, o programa deixa de apostar em diretrizes da FL que são praticamente impossíveis de serem reduzidas ao seu correlato neurofisiológico e genético (Restrição de Ilha, por exemplo) e passa a se concentrar em princípios operativos mínimos – tais como concatenação, recursão, linearização e organização hierárquica (POEPPEL & EMBICK, 2006) – cuja implementação neurológica e codificação genética parecem ser significativamente mais plausíveis. Além disso, a transversalidade interespecífica de alguns desses princípios, que podem eventualmente fazer parte da estrutura neurocognitiva de outros animais (ainda que aplicados à resolução de outras demandas computacionais que não o processamento linguístico), abre caminho para investigações mais profundas sobre os mecanismos neurológicos e genéticos que subjazem à sua realização, permitindo análises comparativas com a espécie humana (incluindo o confronto de grupos de genes) a fim de elucidar aspectos centrais da FL. Como vimos em alguns momentos de nosso trabalho, estudos com modelos animais podem ser produtivos para a Linguística, mesmo quando a distância evolutiva entre as espécies parece muito grande. Análises de

⁹⁵ “If we find evidence for recursion in animals, but in a noncommunicative domain, then we are more likely to pinpoint the mechanisms underlying this ability and the selective pressures that led to it. This discovery, in turn, would open the door to another suite of puzzles: Why did humans, but no other animal, take the power of recursion to create an open-ended and limitless system of communication? Why does our system of recursion operate over a broader range of elements or inputs (e.g., numbers, words) than other animals? One possibility, consistent with current thinking in the cognitive sciences, is that recursion in animals represents a modular system designed for a particular function (e.g., navigation) and impenetrable with respect to other systems. During evolution, the modular and highly domain-specific system of recursion may have become penetrable and domain-general. This opened the way for humans, perhaps uniquely, to apply the power of recursion to other problems. This change from domain-specific to domain-general may have been guided by particular selective pressures, unique to our evolutionary past, or as a consequence (by-product) of other kinds of neural reorganization. Either way, these are testable hypotheses, a refrain that highlights the importance of comparative approaches to the faculty of language.”

competências cognitivas associadas a versões do *FOXP2* em outros animais (principalmente em determinados grupos de aves), por exemplo, têm nos auxiliado a descobrir possíveis funções desse gene na arquitetura da linguagem em humanos.

Pode valer a pena lembrar que os fundamentos da recombinação gênica foram mais bem evidenciados a princípio no estudo das bactérias e da mosca-das-frutas, os mecanismos do desenvolvimento embrionário no estudo do ouriço-do-mar e a dinâmica do impulso nervoso no estudo do axônio gigante da lula. Nos últimos anos, tornou-se padrão explorar a deficiência mental em crianças estudando alelos mutantes específicos em drosófila [...]. Obviamente, nós, humanos, não evoluímos diretamente de nenhuma dessas espécies, mas há semelhanças suficientes em órgãos e funções particulares para tornar esses sistemas-modelo análogos interessantes. Assim, analisar o que um gene como o *Foxp2* pode estar fazendo em outras espécies pode nos dizer muita coisa a respeito da nossa. (PIATTELLI-PALMARINI & URIAGEREKA, 2011: 119)⁹⁶.

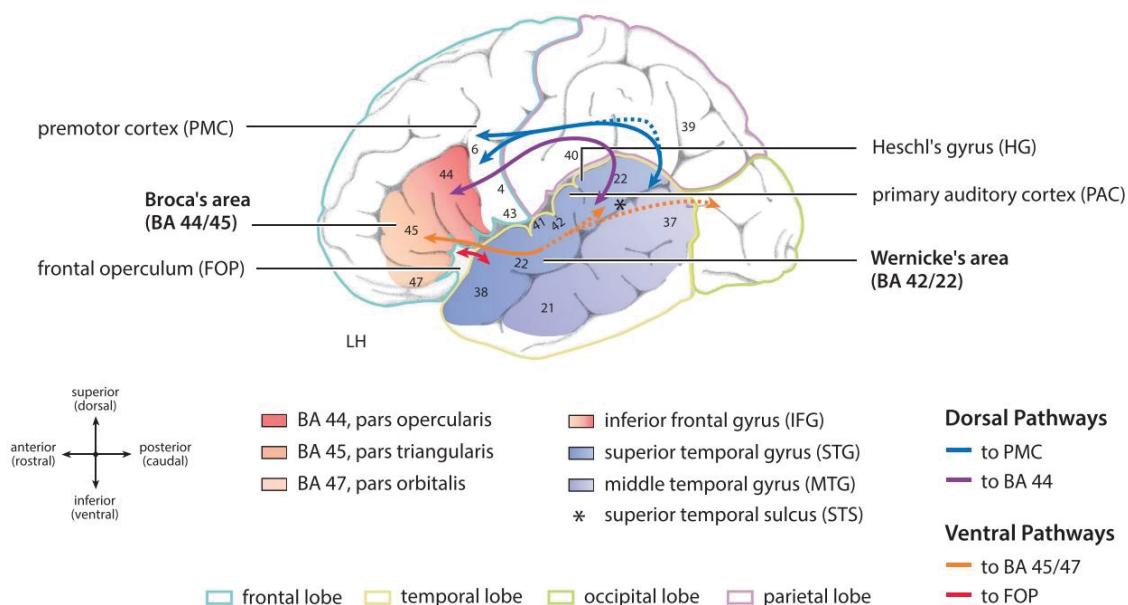
Obviamente, encontrar a etiologia genética das operações mínimas da FL, descobrindo quais genes são responsáveis por sua emergência, continua sendo um imenso desafio – ou, parafraseando o título do texto de Piattelli-Palmarini & Uriagereka (2011) sobre o *FOXP2*, “*a geneticist’s dream, a linguist’s nightmare*.” No entanto, com os crescentes avanços tecnológicos nas áreas de neurociências, genômica e informática, o número de informações acumuladas vem crescendo exponencialmente e algo que parecia impossível há algumas décadas – encontrar no código genético humano o (possível) programa molecular responsável pela faculdade da linguagem em nossa espécie – começa a se tornar factível. De toda forma, desde que a GGT assumiu como um de seus fundamentos a ideia chomskyana de que a FL é uma dotação biológica condicionada geneticamente, assumiu também o compromisso de responder ao problema (1) de modo minimamente adequado – seja isso um pesadelo ou um sonho.

Seja qual for a natureza real da FL (a hipótese do PM é apenas uma entre tantas), decifrá-la é um dos imperativos do nosso tempo e a Linguística tem um papel fundamental na

⁹⁶ “It may repay to be reminded that the essentials of genetic recombination have been best evidenced at first in the study of bacteria and the fruitfly, the mechanisms of embryonic development in the study of the sea urchin and the dynamics of the nervous impulse in the study of the squid’s giant axon. In recent years it has become standard to explore mental retardation in children by studying specific mutants in drosophila [...]. Obviously, we humans did not directly evolve from any of these species, but there are enough commonalities in particular organs and functions to make these model-systems interesting analogs, and thus looking into what a gene like FoxP2 may be doing in other species is very telling.”

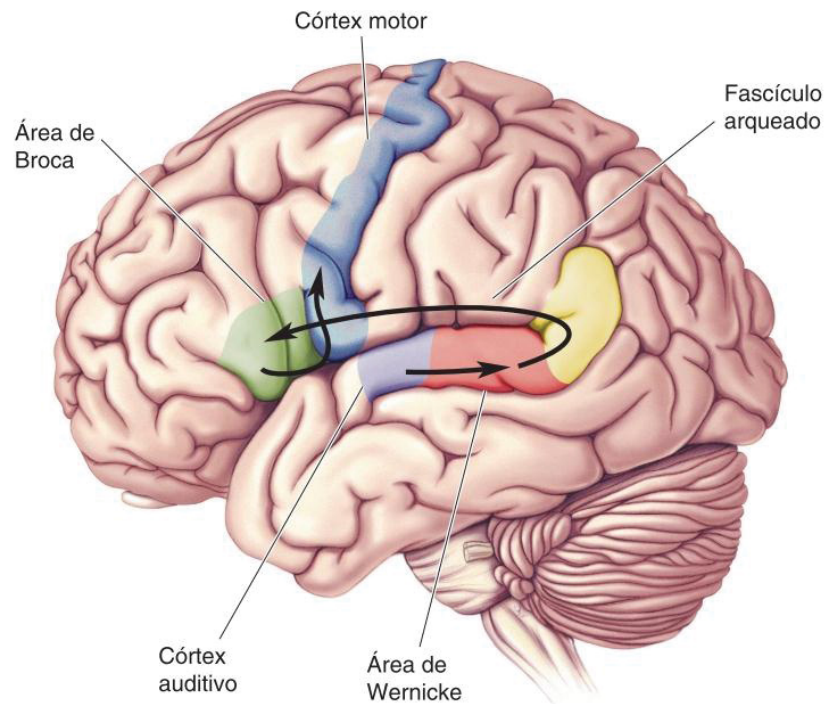
tarefa de construir uma teoria gramatical que integre linguagem, cérebro, genoma e evolução – mesmo porque todas essas dimensões *estão integradas* em cada um de nós. Neste sentido, a pergunta apresentada em (1) deve ser entendida sobretudo como um convite à pesquisa. Longe de oferecermos respostas definitivas para o problema, nosso objetivo foi apresentá-lo como uma diretriz para futuras investigações. Esperamos que o(a) leitor(a) tenha se sentido motivado(a) a enfrentar esse desafio.

APÊNDICE



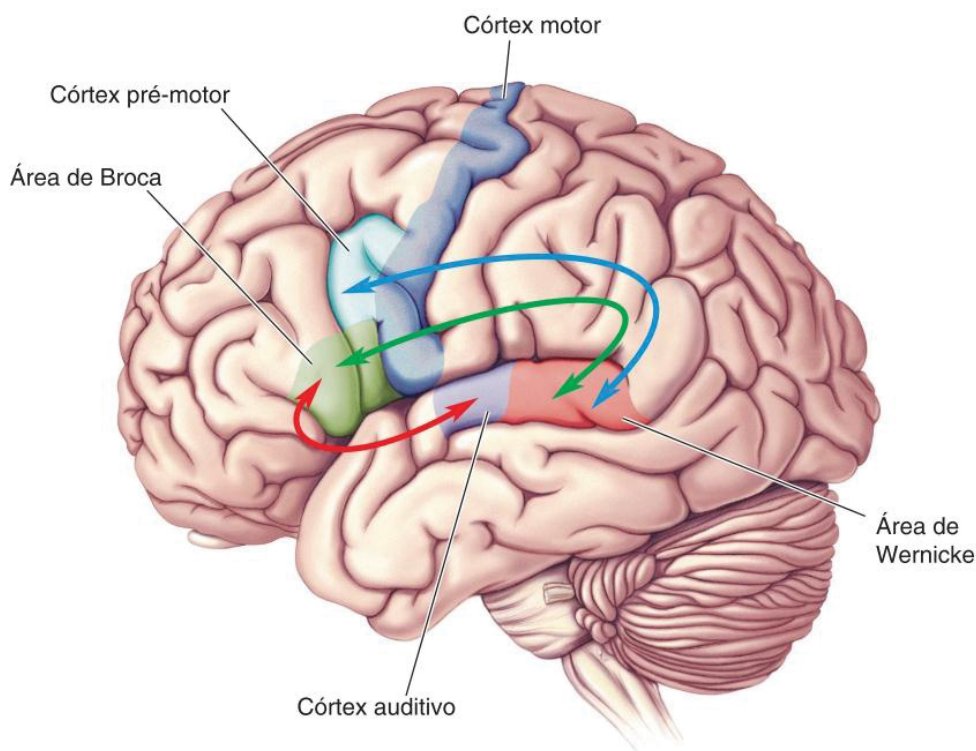
Neuroanatomia da linguagem. Detalhes anatômicos e citoarquitetônicos do hemisfério esquerdo (LH) do cérebro humano. Topo: Os diferentes lobos (frontal, temporal, parietal, occipital) são marcados por bordas coloridas. Os principais giros relevantes para a linguagem (giro frontal inferior (IFG), giro temporal superior (STG), giro temporal médio (MTG)) são codificados por cores; o sulco temporal superior (STS), localizado entre STG e MTG, é marcado por um asterisco. Os números indicam áreas de Brodmann (BA) relevantes para a linguagem que Brodmann (1909) definiu com base em características citoarquitetônicas. Os rótulos de coordenadas (ver canto inferior esquerdo) *superior* (dorsal) – *inferior* (ventral) indicam a posição do giro dentro de um lobo (por exemplo, giro temporal superior) ou dentro de uma BA (por exemplo, BA 44 superior). Os rótulos de coordenadas horizontais *anterior* (rostral) – *posterior* (caudal) indicam a posição dentro de um giro (por exemplo, giro temporal superior anterior). A área de Broca consiste na *pars opercularis* (BA 44) e na *pars triangularis* (BA 45). Localizado anteriormente à área de Broca está a *pars orbitalis* (BA 47). O opérculo frontal (FOP) está localizado ventralmente e medialmente a BA 44, BA 45. O córtex pré-motor (PMC) está localizado em BA 6. A área de Wernicke é definida como BA 42 e BA 22. O córtex auditivo primário (PAC) e giro de Heschl (HG) estão localizados em uma orientação lateral para medial no lobo temporal. Os tratos de fibra da substância branca, ou seja, as vias dorsais e ventrais que conectam as regiões cerebrais relevantes para a linguagem são indicadas por setas codificadas por cores (cf. o canto inferior direito). Extraído e adaptado de FRIEDERICI (2017).

APÊNDICE (Continuação)



Modelo de Wernicke-Geschwind de processamento neurológico da linguagem. Os elementos anatômicos básicos desse sistema são a área de Broca, a área de Wernicke, o fascículo arqueado (um feixe de axônios que conecta essas duas áreas corticais) e o giro angular (destacado em amarelo). O modelo também inclui áreas sensoriais (córtex auditivo) e motoras (córtex motor) envolvidas na recepção e na produção da linguagem. As setas indicam a via de processamento linguístico que é acionada em uma tarefa de recepção e repetição de palavras. Quando os sons que codificam a palavra chegam ao ouvido, o sistema auditivo os processa, gerando sinais neurais que chegam ao córtex auditivo. De acordo com esse modelo, a sequência de sons só é compreendida como uma palavra com significado após o processamento na área de Wernicke. Para que a pessoa possa repetir a palavra, sinais extraídos dela são transmitidos da área de Wernicke para a área de Broca através do fascículo arqueado. Na área de Broca, esses sinais são convertidos em um código que pode ser interpretado pelo sistema de programação motora da fala. Após essa conversão, o código é enviado para as áreas corticais responsáveis por programar e executar os movimentos dos componentes do aparelho fonador – lábios, língua, laringe, cordas vocais – que, em conjunto, permitem a articulação vocal e a realização da fala. Extraído e adaptado de BEAR *et al.* (2017).

APÊNDICE (Continuação)



Modelo de vias paralelas de processamento neurológico da linguagem. Os modelos atuais de processamento neurológico da linguagem se baseiam em múltiplos fluxos de informações que ocorrem em paralelo. O modelo aqui mostrado (BERWICK *et al.*, 2013) inclui duas vias posteriores e uma via anterior. Uma das vias posteriores (em azul) conecta o giro temporal superior (área de Wernicke e córtex auditivo adjacente) ao córtex pré-motor e está envolvida na programação dos movimentos responsáveis pela articulação da fala (incluindo tarefas como a repetição de palavras). A outra via posterior (em verde) conecta o giro temporal superior à área de Broca e, segundo Berwick *et al.* (2013), pode estar envolvida no processamento de estruturas sintáticas complexas (neste modelo, a região BA 44 da área de Broca é responsável pela (re)construção da estrutura de constituintes ao passo que a integração entre a informação sintática e a informação semântica necessária à interpretação da sentença é realizada pela região posterior do giro temporal superior [STG]. Cf. a via destacada em roxo no mapa da pág. 213). O fluxo anterior (em vermelho) processa semanticamente os sons da fala e é corresponsável pela extração de seus significados. Extraído e adaptado de BEAR *et al.* (2017).

REFERÊNCIAS

ADAMS, D. & WASSERSTEIN, M. (2003 [Updated 2020]). Free Sialic Acid Storage Disorders. In: ADAM, M. P.; ARDINGER, H. H.; PAGON, R. A. *et al.* (eds.). *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle.

ADES, C. (1995). A construção da teia geométrica enquanto instinto: primeira parte de um argumento. *Psicologia USP*, v.6, n.1, pp. 145-172.

_____. (2009). Darwin, instinto e mente. *Revista Pesquisa Fapesp*, n.157, pp. 44-45.

ALANO, A.; ALMASHANU, S.; CHIMSKY, J. M.; COSTEAS, P.; BLITZER, M. G.; WULFSBERG, E. A. & COWAN, T. M. (1998). Molecular characterization of a unique patient with epimerase-deficiency galactosaemia. *Journal of inherited metabolic disease*, 21(4), pp. 341-350.

AL-TAHAN, A. Y.; DIVAKARAN, M. P.; KAMBOURIS, M.; BOHLEGA, S.; SALIH, M.; OGUNNIYI, A. & AL-GHANMI, H. (1999). A novel autosomal recessive “Huntington's disease-like” neurodegenerative disorder in a Saudi family. *Saudi medical journal*, 20(1), pp. 85-89.

ALTARESCU, G.; SUN, M.; MOORE, D. F. ; SMITH, J. A.; WIGGS, E. A.; SOLOMON, B. I.; PATRONAS, N. J.; FREI, K. P.; GUPTA, S.; KANESKI, C. R.; QUARRELL, O. W.; SLAUGENHAUPT, S. A.; GOLDIN, E. & SCHIFFMANN, R. (2002). The neurogenetics of mucopolipidosis type IV. *Neurology*, 59(3), pp. 306–313.

AMABIS, J. M. & MARTHO, G. R. (2009). *Biologia das células* (3ª edição.). São Paulo: Moderna.

ANDERSON, S. R. & LIGHTFOOT, D.W. (2000). The human language faculty as an organ. *Annual Review of Physiology*, 62, pp. 697-722.

ANDREW, S. E.; GOLDBERG, Y. P.; KREMER, B.; SQUITIERI, F.; THEILMANN, J.; ZEISLER, J.; TELENUS, H.; ADAM, S.; ALMQUIST, E. & ANVRET, M. (1994). Huntington disease without CAG expansion: phenocopies or errors in assignment?. *American journal of human genetics*, 54(5), pp. 852-863.

- ATKINS, P. W.; JONES, L. L. & LAVERMAN, L. E. (2013). *Chemical principles: the quest for insight* (6^a ed.). New York: W. H. Freeman and Company.
- AZEN, C.; KOCH, R.; FRIEDMAN, E.; WENS, E. & FISHLER, K. (1996). Summary of findings from the United States Collaborative Study of children treated for phenylketonuria. *European journal of pediatrics*, 155, Suppl. 1, S29–S32.
- BADDELEY, A. D.; GATHERCOLE, D. E. & PAPAGNO, C. (1998). The phonological loop as a language learning device. *Psychology Review*. 105, pp. 158-173.
- BAERTLING, F.; RODENBURG, R. J.; SCHAPER, J.; SMEITINK, J. A.; KOOPMAN, W. J.; MAYATEPEK, E.; MORAVA, E. & DISTELMAIER, F. (2014). A guide to diagnosis and treatment of Leigh syndrome. *Journal of neurology, neurosurgery, and psychiatry*, 85(3), pp. 257–265.
- BAKKER, H. D.; DE SONNAVILLE, M. L.; VREKEN, P.; ABELING, N. G.; GROENER, J. E.; KEULEMANS, J. L. & VAN DIGGELEN, O. P. (2001). Human alpha-N-acetylgalactosaminidase (alpha-NAGA) deficiency: no association with neuroaxonal dystrophy? *European journal of human genetics*, 9(2), pp. 91–96.
- BANERJEE, S. & BAXEVANIS, A. D. (2004). Structural analysis of disease-causing mutations in the P-subfamily of forkhead transcription factors. *Proteins*, 54(4), pp. 639–647.
- BARRY, J.; YASIN, I. & BISHOP, D. V. (2007). Heritable risk factors associated with language impairments. *Genes Brain & Behavior*, 6, pp. 66-76.
- BARTLETT, C. W.; FLAX, J. F.; LOGUE, M. W.; VIELAND, V. J.; BASSETT, A. S.; TALLAL, P. & BRZUSTOWICS, L. M. (2002). A major susceptibility locus for specific language impairment is located on 13q21. *American journal of human genetics*, 71, pp. 45-55.
- BATESON, P. P. G. (2014). Behavioural development in the cat. In: TURNER, D. C.; BATESON, P. P. G. (eds.). *The domestic cat: the biology of its behaviour* (3^a edição). Cambridge: Cambridge University Press.
- BEAR, M. F.; CONNORS, B. W.; PARADISO, M.A. (2017). *Neurociências: desvendando o sistema nervoso*. (4^a edição). Tradução de Carla Dalmaz *et al.*. Porto Alegre: Artmed.

BENVENISTE, E. (1952). Comunicação animal e linguagem humana. In: BENVENISTE, E. *Problemas de linguística geral*. Tradução de Glória Novak e Luiza Neri. São Paulo: Editora Nacional, Editora da Universidade de São Paulo, 1976.

BERWICK, R. C. (2011). All you need is Merge: biology, computation, and language from the bottom up. In: DI SCIULLO, A. M. & BOECKX, C. (eds.). *The biolinguistic enterprise: new perspectives on the evolution and nature of the human language faculty*. Oxford: Oxford University Press, pp. 461-491.

BERWICK, R. C.; FRIEDERICI, A. D.; CHOMSKY, N. & BOLHUIS, J. J. (2013). Evolution, brain, and the nature of language. *Trends in cognitive sciences*, 17(2), pp. 89-98.

BICKERTON, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7 (2), pp. 173-221.

BINDU, P. S.; NAGAPPA, M.; BHARATH, R. D. *et al.* (2017). Isolated Sulfite Oxidase Deficiency. In: ADAM, M. P.; ARDINGER, H. H.; PAGON, R. A. *et al.* (eds.). *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle.

BISHOP, D. V. M. (2002a). Motor immaturity and specific speech and language impairment: Evidence for a common genetic basis. [Neuropsychiatric Genetics]. *American journal of medical genetics*, 114, pp. 56-63.

BISHOP, D. V. M. (2002b). The role of genes in the etiology of specific language impairment. *Journal of communication disorders*, 35, pp. 311-328.

BISHOP, D. V. M. & EDMUNDSON, A. (1986). Is otitis media a major cause of specific developmental language disorders?. *British Journal of Disorders of Communication*, 21, pp. 321-338.

BISHOP, D. V. M.; NORTH, T. & DONLAN, C. (1995). Genetic basis of specific language impairment: Evidence from a twin study. *Developmental medicine and child neurology*, 37, pp. 56-71.

BISHOP, D. V. M.; BISHOP, S. J.; BRIGHT, P.; JAMES, C.; DELANEY, T. & TALLAL, P. (1999). Different origin of auditory and phonological processing problems in children with language impairment: evidence from a twin study. *Journal of speech, language and hearing research*, 42, pp. 155-68.

BISHOP, D. V. M.; ADAMS, C. & NORBURY, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: Evidence from 6-year-old twins. *Genes brain & behavior*, 5, pp. 158-169.

BLEY, A.; DENECKE, J.; KOHLSCHÜTTER, A.; SCHÖN, G.; HISCHKE, S.; GUDER, P.; BIERHALS, T.; LAU, H.; HEMPEL, M. & EICHLER, F. S. (2021). The natural history of Canavan disease: 23 new cases and comparison with patients from literature. *Orphanet journal of rare diseases*, 16(1), p. 227.

BOECKX, C. (2011). General preface. In: DI SCIULLO, A. M. & BOECKX, C. (eds.). *The biolinguistic enterprise: new perspectives on the evolution and nature of the human language faculty*. Oxford: Oxford University Press.

BONAGLIA, M. C.; GIORDA, R.; BORGATTI, R.; FELISARI, G.; GAGLIARDI, C.; SELICORNI, A. & ZUFFARDI, O. (2001) Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *American Journal of Human Genetics*, 69(2), pp. 261-268.

BORGES-OSÓRIO, M. R. & ROBINSON, W. M. (2013). *Genética humana* (3^a edição). Porto Alegre: Artmed.

BOSCH, A. M. (2006). Classical galactosaemia revisited. *Journal of inherited metabolic disease*, 29(4), pp. 516-525.

BOWSER, L. E.; YOUNG, M.; WENGER, O. K.; AMMOUS, Z.; BRIGATTI, K. W.; CARSON, V. J.; MOSER, T.; DELINE, J.; AOKI, K.; MORLET, T.; SCOTT, E. M.; PUFFENBERGER, E. G.; ROBINSON, D. L.; HENDRICKSON, C.; SALVIN, J.; GOTTLIEB, S.; HEAPS, A. D.; TIEMEYER, M. & STRAUSS, K. A. (2019). Recessive GM3 synthase deficiency: Natural history, biochemistry, and therapeutic frontier. *Molecular genetics and metabolism*, 126(4), pp. 475–488.

BRANDT, R.; ROHLFING, T.; RYBAK, J.; KROFCZIK, S.; MAYE, A.; WESTERHOFF, M.; HEGE, H. C. & MENZEL, R. (2005). Three-dimensional average-shape atlas of the honeybee brain and its applications. *The Journal of comparative neurology*, 492(1), pp. 1-19.

BROWN, F. H.; McDOUGALL, I. & FLEAGLE, J. G. (2012). Correlation of the KHS Tuff of the Kibish Formation to volcanic ash layers at other sites, and the age of early Homo sapiens (Omo I and Omo II). *Journal of human evolution*, 63(4), pp. 577-585.

- BROWN, P.; GOLDFARB, L. G.; KOVANEN, J.; HALTIA, M.; CATHALA, F.; SULIMA, M.; GIBBS, C. J. & GAJDUSEK, D. C. (1992). Phenotypic characteristics of familial Creutzfeldt-Jakob disease associated with the codon 178Asn PRNP mutation. *Annals of neurology*, 31(3), pp. 282–285.
- BROWN, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- BRUENING, B. (2014). Precede-and-command revisited. *Language*, v.90, n. 2, pp. 342-388.
- BURGESS, P. & SHALLICE, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, 34(4), pp. 263-272.
- BURRACO, A. B. (2012). *Genes y lenguaje, Aspectos ontogenéticos, filogenéticos y cognitivos*. Barcelona: Editorial Reverté S.A.
- CAMPBELL, N. A.; URRY, L. A.; CAIN, M. L.; WASSERMAN, S. A.; MINORSKY, P. V. & ORR, R. B. (2020). *Campbell biology* (Twelfth Edition). New York, NY: Pearson.
- CANALES, J. J. & GRAYBIEL, A. M. (2000). A measure of striatal function predicts motor stereotypy. *Nature neuroscience*, 3(4), pp. 377-383.
- CARNIE, A. (2006). *Syntax: A generative introduction*. Cambridge (MA): Blackwell.
- CARRERA, P.; PIATTI, M.; STENIRRI, S.; GRIMALDI, L. M. E.; MARCHIONI, E.; CURCIO, M.; RIGHETTI, P. G.; FERRARI, M. & GELFI, C. (1999). Genetic heterogeneity in Italian families with familial hemiplegic migraine. *Neurology*, 53, pp. 26-32.
- CASTRO-REBOLLEDO, R.; GIRALDO-PRIETO, M.; HINCAPIÉ-HENAO, L. LOPERA, F. & PINEDA, D. A. (2004). Trastorno específico del desarrollo del lenguaje: una aproximación teórica a su diagnóstico, etiología y manifestaciones clínicas. *Revista de Neurología*. 39, pp. 1173-1181.
- CHENERY, H. J.; COPLAND, D. A. & MURDOCH, B. E. (2002). Complex language functions and subcortical mechanisms: evidence from Huntington's disease and patients with non-thalamic subcortical lesions. *International journal of language & communication disorders*, 37(4), pp. 459-474.

Chimpanzee Sequencing and Analysis Consortium (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, 437 (7055), pp. 69-87.

CHOMSKY, N. (1975). *Reflections on language*. New York: Pantheon.

_____. (1981). *Lectures on government and binding*. Dordrecht: Foris.

_____. (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge: The MIT Press.

_____. (1986). *Knowledge of language: Its nature, origin and use*. New York: Praeger.

_____. (1995). *The minimalist program*. Cambridge: MIT Press.

_____. (2011). Language and other cognitive systems. What is special about language? *Language Learning and Development*, vol.7:4, pp. 263-278.

CHOMSKY, N. & LASNIK, H. (1993). The theory of principles and parameters. In: JACOBS, J.; STECHOW, A. STERNEFELD, W. & VENNEMANN, T. (eds.). *Syntax: an international handbook of contemporary research*. Berlin: Walter de Gruyter, pp. 506-569.

CLARK, M. A.; DOUGLAS, M. & CHOI, J. (2018). *Biology 2e*. Houston: OpenStax.

CLARKE, L. A. (2002 [Updated 2021]). Mucopolysaccharidosis Type I. In: ADAM, M. P.; ARDINGER, H. H.; PAGON, R. A. *et al.* (eds.). *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle.

COELHO, S.; ALBUQUERQUE, C. P. & SIMÕES, M. R. (2013). Distúrbio Específico de Linguagem: caracterização neuropsicológica. *Paidéia*, 23(54), pp. 31-41.

COMINGS, D. E.; WU, S.; CHIU, C.; MUHLEMAN, D. & SVERD, J. (1996). Studies of the c-Harvey-Ras gene in psychiatric disorders. *Psychiatry research*, 63(1), pp. 25-32.

CONTI-RAMSDEN, G.; SIMKIN, Z. & PICKLES, A. (2006). Estimating familial loading in SLI: A comparison of direct assessment versus parental interview. *Journal of Speech, Language and Hearing Research*, 49, pp. 88-101.

COOKE, S. J.; CROSSIN, G. T.; HINCH, S. G. (2011). Pacific salmon migration: completing the cycle. In: FARREL, A. P. (ed.). *Encyclopedia of fish physiology: from genome to environment*. San Diego: Academic Press, v.3. pp. 1945-1952.

COPE, N.; HAROLD, D.; HILL, G.; MOSKVINA, V.; STEVENSON, J.; HOLMANS, P.; OWEN, M. J.; O'DONOVAN, M. C. & WILLIAMS, J. (2005). Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *American journal of human genetics*, 76(4), pp. 581-591.

CORBALLIS, M. C. (2004). FOXP2 and the mirror system. *Trends in cognitive sciences*, 8(3), pp. 95-96.

CRESTANI, A. H.; OLIVEIRA, L. D.; VENDRUSCOLO, J. F. & RAMOS-SOUZA, A. P. (2013). Distúrbio específico de linguagem: a relevância do diagnóstico inicial. *Revista CEFAC*, 15(1), pp. 228-37.

De CLEMENTE, V.; VITIELLO, G.; IMPERATI, F.; ROMANO, A.; PARENTE, I.; ROSA, M.; PASCARELLA, A.; PARENTI, G. & Del GIUDICE, E. (2013). Smith Lemli-Opitz syndrome: a contribution to the delineation of a cognitive/behavioral phenotype. *Minerva pediatrica*, 65(1), pp. 61-69.

De FELICE, S.; ROMANI, C.; GEBERHIWOT, T.; MacDONALD, A. & PALERMO, L. (2018). Language processing and executive functions in early treated adults with phenylketonuria (PKU). *Cognitive neuropsychology*, 35(3-4), pp. 148-170.

De KOVEL, C. G. F.; HOL, F. A.; HEISTER, J. G. A. M.; WILLEMEN, J. J. H. T.; SANDKUIJL, L. A.; FRANKE, B. & PADBERG, G. W. (2004). Genomewide scan identifies susceptibility locus for dyslexia on Xq27 in an extended Dutch family. *Journal of medical genetics*, 41, pp. 652-657.

DEACON, T. W. (2000). Evolutionary perspectives on language and brain plasticity. *Journal of communication disorders*, 33(4), pp. 273-291.

DEO, P.N. & DESHMUKH, R. (2019). Oral microbiome: Unveiling the fundamentals. *Journal of oral and maxillofacial pathology*, 23(1), pp. 122-128.

DESMOND, J. E. & FIEZ, J. A. (1998). Neuroimaging studies of the cerebellum: language, learning and memory. *Trends in cognitive sciences*, 2(9), pp. 355-362.

DeTHORNE, L.; HART, S.; PETRILL, S.; DEATER-DECKARD, K.; THOMPSON, L. A.; SCHATSCHNEIDER, C. *et al.* (2006). Children's history of speech-language difficulties:

Genetic influences and associations with reading-related measures. *Journal of Speech, Language and hearing research*, 49, pp. 1280–1293.

DeTHORNE, L. S.; JOHNSON, B. W.; LOEB, J. W. (2005). A closer look at MLU: what does it really measure? *Clinical Linguistics & Phonetics*, 19(8), pp. 635-648.

DORLAND, L.; DURAN, M.; HOEFNAGELS, F. E.; BREG, J. N.; FABERY de JONGE, H.; CRANSBERG, K.; VAN SPRANG, F. J. & VAN DIGGELEN, O. P. (1988). Beta-mannosidosis in two brothers with hearing loss. *Journal of inherited metabolic disease*, 11, Suppl. 2, pp. 255–258.

DUARTE, M. E. L. (1995). *A perda do princípio “Evite Pronome” no Português Brasileiro*. Tese de Doutorado. Campinas: UNICAMP.

EMBICK, D.; MARANTZ, A.; MIYASHITA, Y; O'NEIL, W. & SAKAI, K. L. (2000). A syntactic specialization for Broca's area. *Proceedings of the National Academy of Sciences of the United States of America*, 97(11), pp. 6150-6154.

ENARD, W.; PRZEWORSKI, M.; FISHER, S. E.; LAI, C. S.; WIEBE, V.; KITANO, T.; MONACO, A. P. & PÄÄBO, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature*, 418(6900), pp. 869-872.

FAGERHEIM, T.; RAEYMAEKERS, P.; TØNNESSEN, F. E.; PEDERSEN, M.; TRANEBJAERG, L. & LUBS, H. A. (1999). A new gene (DYX3) for dyslexia is located on chromosome 2. *Journal of medical genetics*, 36(9), pp. 664-669.

FALCARO, M.; PICKLES, A.; NEWBURY, D. F.; ADDIS, L.; BANFIELD, E.; FISHER, S. E.; MONACO, A. P.; SIMKIN, Z.; CONTI-RAMSDEN, G. & SLI Consortium. (2008). Genetic and phenotypic effects of phonological short-term memory and grammatical morphology in specific language impairment. *Genes, brain, and behavior*, 7(4), pp. 393-402.

FERLAND, R. J.; CHERRY, T. J.; PREWARE, P. O ; MORRISEY, E. E. & WALSH, C. A. (2003). Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain. *The Journal of comparative neurology*, 460(2), pp. 266-279.

FIFER, W. P. & MOON, C. (1989). Auditory experience in the fetus. In: SMOTHERMAN, W. P. & ROBINSON, S. R. (eds.). *Behavior of the fetus*. Caldwell: Telford Press.

FINSTERER, J. (2008). Leigh and Leigh-like syndrome in children and adults. *Pediatric neurology*, 39(4), pp. 223–235.

FISHER, S. E. (2019). Human genetics: The evolving story of FOXP2. *Current biology: CB*, 29(2), R65-R67.

FISHER, S. E. & DEFRIES, J. C. (2002). Developmental dyslexia: genetic dissection of a complex cognitive trait. *Nature reviews. Neuroscience*, 3(10), pp. 767-780.

FISHER, S. E. & FRANCKS, C. (2006). Genes, cognition and dyslexia: learning to read the genome. *Trends in cognitive sciences*, 10(6), pp. 250-257.

FISHER, S. E.; FRANCKS, C.; MARLOW, A. J.; MACPHIE, I. L.; NEWBURY, D. F.; CARDON, L. R.; ISHIKAWA-BRUSH, Y.; RICHARDSON, A. J.; TALCOTT, J. B.; GAYÁN, J.; OLSON, R. K.; PENNINGTON, B. F.; SMITH, S. D.; DEFRIES, J. C.; STEIN, J. F. & MONACO, A. P. (2002). Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nature genetics*, 30(1), pp. 86-91.

FISHER, S. E.; LAI, C. S. & MONACO, A. P. (2003). Deciphering the genetic basis of speech and language disorders. *Annual Review of Neuroscience*. 26, pp. 57-80.

FISHER, S. E.; MARLOW, A. J.; LAMB, J.; MAESTRINI, E.; WILLIAMS, D. F.; RICHARDSON, A. J.; WEEKS, D. E.; STEIN, J. F. & MONACO, A. P. (1999). A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. *American journal of human genetics*, 64(1), pp. 146-156.

FISHER, S. E.; VARGHA-KHADEM, F.; WATKINS, K. E.; MONACO, A. P. & PEMBREY, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nature genetics*, 18(2), pp. 168–170.

FLETCHER, P. (1990). Speech and language defects. *Nature*, 346, p. 226.

FRANCKS, C.; PARACCHINI, S.; SMITH, S. D.; RICHARDSON, A. J.; SCERRI, T. S.; CARDON, L. R.; MARLOW, A. J.; MACPHIE, I. L.; WALTER, J.; PENNINGTON, B. F.; FISHER, S. E.; OLSON, R. K.; DEFRIES, J. C.; STEIN, J. F. & MONACO, A. P. (2004). A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. *American journal of human genetics*, 75(6), pp. 1046-1058.

- FREGE, G. (1978 [1892]). *Lógica e filosofia da linguagem*. São Paulo: Cultrix Editora da Universidade de São Paulo.
- FRESNEDA, M. D. & MENDOZA, E. (2005). Trastorno específico del lenguaje: Concepto, clasificaciones y criterios de identificación. *Revista de Neurología*, 41 (Supl.1): S51-S56.
- FRIEDERICI, A. D. (2017). *Language in our brain. The origins of a uniquely human capacity*. Cambridge: MIT Press.
- FROSTER, U.; SCHULTE-KÖRNE, G.; HEBEBRAND, J. & REMSCHMIDT, H. (1993). Cosegregation of balanced translocation (1,2) with retarded speech development and dyslexia. *Lancet*, 342, pp. 178-179.
- GALLISTEL, C. R. (2010). Learning organs. In: BRICMONT, J. & FRANCK, J. (eds.). *Chomsky notebook*. New York: Columbia University Press.
- GARDNER, R. A. & GARDNER, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, 165 (3894), pp. 664-672.
- GAYÁN, J.; SMITH, S. D.; CHERNY, S. S.; CARDON, L. R.; FULKER, D. W.; BROWER, A. M.; OLSON, R. K.; PENNINGTON, B. F. & DEFRIES, J. C. (1999). Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *American journal of human genetics*, 64(1), pp. 157–164.
- GIBSON, K. M.; WAPPNER, R. S.; JOOSTE, S.; ERASMUS, E.; MIENIE, L. J.; GERLO, E. DESPRECHINS, B. & De MEIRLEIR, L. (1998). Variable clinical presentation in three patients with 3-methylglutaconyl-coenzyme A hydratase deficiency. *Journal of inherited metabolic disease*, 21(6), pp. 631–638.
- GLEITMAN, L. R. & NEWPORT, E. L. (1995). The invention of language by children: environmental and biological influences on the acquisition of language. In: GLEITMAN, L. R. & LIBERMAN, M. (eds.). *An invitation to cognitive science*. Vol. 1: Language. Cambridge: The MIT Press.
- GONÇALVES, R. (2006). *A expressão do FOXP2. Uma introdução ao estudo das relações entre genes e linguagem*. Porto Alegre: Universidade Federal do Rio Grande do Sul.
- GOPNIK, M. (1990). Feature-blind grammar and dysphasia. *Nature*, 344(6268), p. 715.

- GOPNIK, M. & CRAGO, M. B. (1991). Familial aggregation of a developmental language disorder. *Cognition*, 39, pp. 1-50.
- GOPNIK, M.; DALALAKIS, J.; FUKUDA, S. E.; FUKUDA, S. & KEHAYIA, E. (1996). Genetic language impairment: unruly grammars. *Proceedings of the British Academy*, 88, pp. 223-249.
- GRIFFITHS, A. J. F.; WESSLER, S. R.; CARROL, S. B.; DOEBLEY, J. (2013). *Introdução à genética* (10^a edição). Tradução de Idília Vanzelotti. Rio de Janeiro: Guanabara Koogan.
- GRIGORENKO, E. L.; WOOD, F. B.; MEYER, M. S. & PAULS, D. L. (2000). Chromosome 6p influences on different dyslexia-related cognitive processes: further confirmation. *American journal of human genetics*, 66(2), pp. 715–723.
- GRODZINSKY, Y. (2000). The neurology of syntax: language use without Broca's area. *The Behavioral and brain sciences*, 23(1), pp. 1-71.
- GROLLA, E. (2013). A aquisição do princípio C da teoria de ligação em português brasileiro: questões metodológicas. *Revista de estudos linguísticos*, v.21, n.2. pp. 9-34.
- GRÜTER, C. & FARINA, W. M. (2009). The honeybee waggle dance: can we follow the steps? *Trends in ecology & evolution*, 24(5), pp. 242-247.
- GUASTI, M. T. (2002). *Language acquisition. The growth of grammar*. Cambridge: The MIT Press.
- GUEVARA-CAMPOS, J.; GONZÁLEZ-GUEVARA, L.; GUEVARA-GONZÁLEZ, J. & CAULI, O. Guevara-Campos, J. (2019). First Case Report of Primary Carnitine Deficiency Manifested as Intellectual Disability and Autism Spectrum Disorder. *Brain sciences*, 9(6), p. 137.
- GUIMARÃES, M. (2017). *Os fundamentos da teoria linguística de Chomsky*. Petrópolis: Vozes.
- HAEGEMAN, L. (1994). *Introduction to government and binding theory*. Cambridge, MA: Blackwell.
- HAESLER, S.; WADA, K.; NSHDEJAN, A.; MORRISEY, E. E.; LINTS, T.; JARVIS, E. D. & SCHARFF, C. (2004). FoxP2 expression in avian vocal learners and non-learners. *The*

Journal of neuroscience : the official journal of the Society for Neuroscience, 24(13), pp. 3164-3175.

HAHN, K. A.; SALOMONS, G. S.; TACKELS-HORNE, D.; WOOD, T. C.; TAYLOR, H. A.; SCHROER, R. J.; LUBS, H. A.; JAKOBS, C.; OLSON, R. L.; HOLDEN, K. R.; STEVENSON, R. E. & SCHWARTZ, C. E. (2002). X-linked mental retardation with seizures and carrier manifestations is caused by a mutation in the creatine-transporter gene (SLC6A8) located in Xq28. *American journal of human genetics*, 70(5), pp. 1349–1356.

HANNULA-JOUPPI, K.; KAMINEN-AHOLA, N.; TAIPALE, M.; EKLUND, R.; NOPOLA-HEMMI, J.; KÄÄRIÄINEN, H. & KERE, J. (2005). The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. *PLoS genetics*, 1(4), e50.

HAYNES, C. & NAIDOO, S. (1991). *Children with specific speech and language impairment*. London: Mac Keith Press.

HAUSER, M. D.; CHOMSKY, N. & FITCH, W. T. (2002). The language faculty: what is it, who has it, and how did it evolve? *Science*, n. 298, pp. 1569-79.

HEDIGER, H. K. P. (1981). The Clever Hans phenomenon from an animal psychologist's point of view. In: SEBEOK, T. A. & ROSENTHAL, R. (eds.). *The Clever Hans phenomenon: communication with horses, whales, apes, and people*. New York: The New York Academy of Sciences. Annals of the New York Academy of Sciences, vol. 364.

HENNERMANN, J. B. (2006). Clinical variability in glycine encephalopathy. *Future Neurology*, 1, pp. 621-30.

HOFMANN, H. A. (2003). Functional genomics of neural and behavioral plasticity. *Journal of Neurobiology*, 54(1), pp. 272-282.

HOLMBERG, A. & ROBERTS, I. (2009). Introduction: parameters in minimalist theory. In: BIBERAUER, T. et al. *Parametric variation: Null subjects in minimalist theory*. Cambridge, MA: Cambridge University Press.

HOOVER-FONG, J. E.; SHAH, S.; VAN HOVE, J. L.; APPLGARTH, D.; TOONE, J. & HAMOSH, A. (2004). Natural history of nonketotic hyperglycinemia in 65 patients. *Neurology*, 63(10), pp. 1847–1853.

HSIUNG, G. Y.; KAPLAN, B. J.; PETRYSHEN, T. L.; LU, S. & FIELD, L. L. (2004). A dyslexia susceptibility locus (DYX7) linked to dopamine D4 receptor (DRD4) region on chromosome 11p15.5. *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*, 125B(1), pp. 112-119.

HURST, J. A.; BARAITSER, M.; AUGER, E.; GRAHAN, F. & NORELL, S. (1990). An extended family with a dominantly inherited speech disorder. *Developmental medicine and child neurology*, 32(4), pp. 352-355.

IJLST, L.; LOUPATTY, F. J.; RUITER, J. P.; DURAN, M.; LEHNERT, W. & WANDERS, R. J. (2002). 3-Methylglutaconic aciduria type I is caused by mutations in AUH. *American journal of human genetics*, 71(6), pp. 1463–1466.

JACOB, P. (2010). The scope and limits of Chomsky's naturalism. In: BRICMONT, J. & FRANCK, J. (eds.). *Chomsky notebook*. New York: Columbia University Press.

JANZEN, D. & NGUYEN, M. (2010). Beyond executive function: non-executive cognitive abilities in individuals with PKU. *Molecular genetics and metabolism*, 99 Suppl. 1: S47–S51.

JAROFF, L. (1989). The Gene Hunt. *Time Magazine*, Mar. 1989, pp. 62-67.

JARVIS, E. D. (2004). Learned birdsong and the neurobiology of human language. *Ann. NY Acad. Sci.* 1016, pp. 749-777.

JERNE, N. K. (1985). The generative grammar of the immune system. Nobel lecture, 8 December 1984. *Bioscience reports*, 5(6), pp. 439-451.

JEZELA-STANEK, A.; CIARA, E. & STEPIEN, K. M. (2020). Neuropathophysiology, Genetic Profile, and Clinical Manifestation of Mucopolysaccharidosis IV-A Review and Case Series. *International journal of molecular sciences*, 21(12), p. 4564.

KAMINEN, N.; HANNULA-JOUPPI, K.; KESTILÄ, M.; LAHERMO, P.; MULLER, K.; KAARANEN, M.; MYLLYLUOMA, B.; VOUTILAINEN, A.; LYYTINEN, H.; NOPOLA-HEMMI, J. & KERE, J. (2003). A genome scan for developmental dyslexia confirms linkage to chromosome 2p11 and suggests a new locus on 7q32. *Journal of medical genetics*, 40(5), pp. 340-345.

KENEDY, E. (2013). *Curso básico de linguística gerativa*. São Paulo: Contexto.

- KOENEMAN, O. & ZEIJLSTRA, H. (2017). *Introducing syntax*. Cambridge: Cambridge University Press.
- KOHN, H.; MANOWITZ, P.; MILLER, M. & KLING. (1988). Neuropsychological deficits in obligatory heterozygotes for metachromatic leukodystrophy. *Human genetics*, 79(1), pp. 8–12.
- KRAUSE, J.; LALUEZA-FOX, C.; ORLANDO, L.; ENARD, W.; GREEN, R. E.; BURBANO, H. A.; HUBLIN, J. J.; HÄNNI, C.; FORTEA, J.; DE LA RASILLA, M.; BERTRANPETIT, J.; ROSAS, A. & PÄÄBO, S. (2007). The derived FOXP2 variant of modern humans was shared with Neandertals. *Current biology: CB*, 17(21), pp. 1908-1912.
- KWASNICKA-CRAWFORD, D. A.; CARSON, A. R.; ROBERTS, W.; SUMMERS, A. M.; REHNSTROM, K.; JARVELA, I. & SCHERER, S. W. (2005). Characterization of a novel cation transporter ATPase gene (ATP13A4) interrupted by 3q25-q29 inversion in an individual with language delay. *Genomics*, 86, pp. 182-194.
- KYLLERMAN, M.; SKJELDAL, O. H.; LUNDBERG, M.; HOLME, I.; JELLUM, E.; von DÖBELN, U.; FOSSEN, A. & CARLSSON, G. (1994). Dystonia and dyskinesia in glutaric aciduria type I: clinical heterogeneity and therapeutic considerations. *Movement disorders : official journal of the Movement Disorder Society*, 9(1), pp. 22-330.
- KYLLERMAN, M.; SKJELDAL, O. H.; CHRISTENSEN, E.; HAGBERG, G.; HOLME, E.; LÖNNQUIST, T.; SKOV, L.; ROTWELT, T. & von DÖBELN, U. (2004). Long-term follow-up, neurological outcome and survival rate in 28 Nordic patients with glutaric aciduria type 1. *European journal of paediatric neurology: official journal of the European Paediatric Neurology Society*, 8(3), pp. 121-129.
- LAHEY, M. & EDWARDS, J. (1995). Specific language impairment: Preliminary investigation of factors associated with family history and patterns of language performance. *Journal of Speech and Hearing Research*, 38, pp. 643-657.
- LAI, C. S.; FISHER, S. E.; HURST, J. A.; LEVY, E. R.; HODGSON, S.; FOX, M.; JEREMIAH, S.; POVEY, S.; JAMISON, D. C.; GREEN, E. D.; VARGHA-KHADEM, F. & MONACO, A. P. (2000). The SPCH1 region on human 7q31: genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *American journal of human genetics*, 67(2), pp. 357-368.

LAI, C. S.; FISHER, S. E.; HURST, J. A.; VARGHA-KHADEM, F. & MONACO, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413(6855), pp. 519-523.

LANDAU, D.; COHEN, D.; SHALEV, H.; PINSK, V.; YERUSHALMI, B.; ZEIGLER, M.; BIRK, O.S. (2004). A novel mutation in the SLC17A5 gene causing both severe and mild phenotypes of free sialic acid storage disease in one inbred Bedouin kindred. *Molecular genetics and metabolism*, 82(2), pp. 167-72.

LANDER, E. S.; LINTON, L. M.; BIRREN, B.; NUSBAUM, C.; ZODY, M. C.; BALDWIN, J.; DEVON, K.; DEWAR, K.; DOYLE, M.; FITZHUGH, W.; FUNKE, R.; GAGE, D.; HARRIS, K.; HEAFORD, A.; HOWLAND, J.; KANN, L.; LEHOCZKY, J.; LEVINE, R.; MCEWAN, P.; MCKERNAN, K. *et al.* International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), pp. 860-921.

LEHMANN, O. J.; SOWDEN, J. C.; CARLSSON, P.; JORDAN, T. & BHATTACHARYA, S. S. (2003). Fox's in development and disease. *Trends in genetics: TIG*, 19(6), pp. 339-344.

LEONARD, L. (2014). *Children with specific language impairment* (Second Edition). Cambridge, MA: MIT Press.

LEWIS, B. & THOMPSON, L. (1992). A study of developmental speech and language disorders in twins. *Journal of speech and hearing research*, 35, pp. 1086–1094.

LIEBERMAN, P. (2002). On the nature and evolution of the neural bases of human language. *American journal of physical anthropology*, Suppl 35, pp. 36-62.

LIEBERMAN, P. (1996). Some biological constraints on the analysis of prosody. In: MORGAN, J. L. & DEMUTH, K. (eds.). *Signal to syntax: bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

LIÉGOIS, F.; BALDEWEG, T.; CONNELLY, A.; GADIAN, D. G.; MISHKIN, M. & VARGHA-KHADEM, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nature neuroscience*, 6(11), pp. 1230-1237.

LIGHTFOOT, D. (1989). The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Sciences*, 12(2), pp. 321-334.

- LINDGREN, K.; FOLSTEIN, S.; TOMBLIN, J. B. & TAGER-FLUSBERG, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research*, 2, pp. 22-38.
- LOBATO, L. (1986). *Sintaxe do português: da teoria padrão à teoria da regência e ligação*. Belo Horizonte: Vigília.
- LOCKE, J. (1999 [1690]). *An essay concerning human understanding*. MANIS, J. (ed.). Electronic Classics Series, The Pennsylvania State University, Hazleton, PA.
- LOPES S. & ROSSO. S. (2010). *Biologia: volume 2*. São Paulo: Saraiva.
- LUST, B. (2006). *Child language: Acquisition and growth*. Cambridge: Cambridge University Press.
- LUST, B.; EISELE, J. & MAZUKA, R. (1992). The binding theory module: Evidence from first language acquisition for Principle C. *Language*, v.68, n.2, pp. 333-358.
- MACERATESI, P.; DAUDE, N.; DALLAPICCOLA, B.; NOVELLI, G.; ALLEN, R.; OKANO, Y. & REICHARDT, J. (1998) Human UDP-galactose 4' epimerase (GALE) gene and identification of five missense mutations in patients with epimerase-deficiency galactosemia. *Molecular genetics and metabolism*, 63(1), pp. 26-30.
- MARCÃO, A. M.; WIEST, R.; SCHINDLER, K.; WIESMANN, U.; WEIS, J.; SCHROTH, G.; MIRANDA, M. C.; STURZENEGGER, M. & GIESELMANN, V. (2005). Adult onset metachromatic leukodystrophy without electroclinical peripheral nervous system involvement: a new mutation in the ARSA gene. *Archives of neurology*, 62(2), pp. 309-313.
- MARCUS, G. F. (2004). *The birth of the mind. How a tiny number of genes creates the complexities of human thought*. New York: Basic Books.
- MARCUS, G. F. & FISHER, S. E. (2003). FOXP2 in focus: what can genes tell us about speech and language?. *Trends in cognitive sciences*, 7(6), pp. 257-262.
- MARGOLIS, R. L.; O'HEARN, E.; ROSENBLATT, A.; WILLOUR, V.; HOLMES, S. E.; FRANZ, M. L.; CALLAHAN, C.; HWANG, H. S.; TRONCOSO, J. C. & ROSS, C. A. (2001). A disorder similar to Huntington's disease is associated with a novel CAG repeat expansion. *Annals of neurology*, 50(3), pp. 373-380.

MARLOW, A. J.; FISHER, S. E.; FRANCK, C.; MACPHIE, I. L.; CHERNY, S. S.; RICHARDSON, A. J.; TALCOTT, J. B.; STEIN, J. F.; MONACO, A. P. & CARDON, L. R. (2003). Use of multivariate linkage analysis for dissection of a complex cognitive trait. *American journal of human genetics*, 72(3), pp. 561-570.

MARQUES, S. F. & LIMONGI, S. C. (2011). Mean length utterance (MLU) as a measure of language development of children with Down syndrome. *Jornal da Sociedade Brasileira de Fonoaudiologia*, 23(2), pp. 152-157.

MATALON, R.; DELGADO, L.; MICHALS-MATALON, K. (1999 [Updated 2018]). Canavan Disease. In: ADAM, M. P.; ARDINGER, H. H.; PAGON, R. A. *et al.*, (eds). *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle.

McCRACKEN, J. T.; SMALLEY, S. L.; MCGOUGH, J. J.; CRAWFORD, L.; DEL'HOMME, M.; CANTOR, R. M.; LIU, A. & NELSON, S. F. (2000). Evidence for linkage of a tandem duplication polymorphism upstream of the dopamine D4 receptor gene (DRD4) with attention deficit hyperactivity disorder (ADHD). *Molecular psychiatry*, 5(5), pp. 531-536.

McGINNIS, W.; GARBER, R. L.; WIRZ, J.; KUROIWA, A. & GEHRING, W. J. (1984). A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell*, n.37, pp. 403-408.

MENG, H., SMITH, S. D., HAGER, K., HELD, M., LIU, J., OLSON, R. K., PENNINGTON, B. F., DEFRIES, J. C., GELERTER, J., O'REILLY-POL, T., SOMLO, S., SKUDLARSKI, P., SHAYWITZ, S. E., SHAYWITZ, B. A., MARCHIONE, K., WANG, Y., PARAMASIVAM, M., LOTURCO, J. J., PAGE, G. P., & GRUEN, J. R. (2005). DCDC2 is associated with reading disability and modulates neuronal development in the brain. *Proceedings of the National Academy of Sciences of the United States of America*, 102(47), pp. 17053-17058.

MIRZOYAN, Z.; SOLLAZZO, M.; ALLOCA, M.; VALENZA, A. M.; GRIFONI, D. & BELLOSTA, P. (2019). *Drosophila melanogaster*: A Model Organism to Study Cancer. *Frontiers in genetics*, v. 10, article 51, pp. 1-16.

MIXSON, A. J.; PARRILLA, R.; RANSOM, S. C.; WIGGS, E. A.; McCLASKEY J. H.; HAUSER, P. & WEINTRAUB, B. D. (1992). Correlations of language abnormalities with

localization of mutations in the beta-thyroid hormone receptor in 13 kindreds with generalized resistance to thyroid hormone: identification of four new mutations. *The Journal of Clinical Endocrinology and Metabolism*, 75(4), pp. 1039-1045.

MOBERGET, T. & IVRY, R. B. (2016). Cerebellar contributions to motor control and language comprehension: searching for common computational principles. *Annals of the New York Academy of Sciences*, 1369(1), pp. 154-171.

MONACO, A. P. (2007). Multivariate linkage analysis of specific language impairment (SLI). *Annals of Human Genetics*, 71, pp. 660-673.

MONTEIRO, P. & FENG, G. (2017). SHANK proteins: roles at the synapse and in autism spectrum disorder. *Nature reviews. Neuroscience*, 18(3), pp. 147-157.

MORENO, A. M.; DE SOUZA, D. d. G. & REINHARD, J. (2012). A comparative study of relational learning capacity in honeybees (*Apis mellifera*) and stingless bees (*Melipona rufiventris*). *PLOS ONE* 7(12): e51467, pp. 1-7.

MORRIS, D. W.; ROBINSON, L.; TURIC, D.; DUKE, M.; WEBB, V.; MILHAM, C.; HOPKIN, E.; POUND, K.; FERNANDO, S.; EASTON, M.; HAMSHIRE, M.; WILLIAMS, N.; MCGUFFIN, P.; STEVENSON, J.; KRAWCZAK, M.; OWEN, M. J.; O'DONOVAN, M. C. & WILLIAMS, J. (2000). Family-based association mapping provides evidence for a gene for reading disability on chromosome 15q. *Human molecular genetics*, 9, pp. 843-848.

MUNTONI, F. & VOIT, T. (2004). The congenital muscular dystrophies in 2004: a century of exciting progress. *Neuromuscular disorders: NMD*, 14(10), pp. 635-649.

MURDOCH, B. E.; OZANNE, A. E. & CROSS, J. A. (1990). Acquired childhood speech disorders: dysarthria and dyspraxia. In: MURDOCH, B.E. (ed.). *Acquired neurological speech/language disorders in childhood*. London: Taylor & Francis.

NABOKOV, V. (2015). *Lições de literatura*. Tradução de Jorio Dauster. São Paulo: Três estrelas.

NEILS, J. & ARAM, D. (1986). Family history of children with developmental language disorders. *Perceptual and Motor Skills*, 63, pp. 655-658.

NELSON, C. D.; WAGGONER, D. D.; DONNELL, G. N.; TUERCK, J. M. & BUIST, N. R. (1991). Verbal dyspraxia in treated galactosemia. *Pediatrics*, 88(2), pp. 346-350.

NEWBURY, D. F.; BONORA, E.; LAMB, J. A.; FIAHER, S. E.; LAI, C. S.; BAIRD, G.; JANNOUN, L.; SLONIMS, V.; STOTT, C. M.; MERRICKS, M. J.; BOLTON, P. F.; BAILEY, A. J.; MONACO, A. P. & International Molecular Genetic Study of Autism Consortium (2002). FOXP2 is not a major susceptibility gene for autism or specific language impairment. *American journal of human genetics*, 70(5), pp. 1318–1327.

NEWBURY, D. F. & MONACO, A. P. (2010). Genetic advances in the study of speech and language disorders. *Neuron*, 68(2), pp. 309-320.

NEWBURY, D. F.; WINCHESTER, L.; ADDIS, L.; PARACCHINI, S.; BUCKINGHAM, L. L.; CLARK, A.; COHEN, W.; COWIE, H.; DWORZYNSKI, K.; EVERITT, A.; GOODYER, I. M.; HENNESSY, E.; KINDLEY, A. D.; MILLER, L. L.; NASIR, J.; O'HARE, A. *et al.* (2009). CMIP and ATP2C2 modulate phonological short-term memory in language impairment. *American journal of human genetics*, 85(2), pp. 264-272.

NOVELLI, G. & REICHARDT, J. K. (2000). Molecular basis of disorders of human galactose metabolism: past, present, and future. *Molecular genetics and metabolism*, 71(1-2), pp. 62-65.

NUSSBAUM, R. L.; McINNES, R. R. & WILLARD, H. F. (2016). *Thompson & Thompson: genética médica* (8ª edição). Tradução de Ana Julia Perrotti-Garcia *et al.*. Rio de Janeiro: Elsevier.

NYHAM, W. L. & HOFFMANN, G. F. (2020). *Atlas of inherited metabolic diseases* (Fourth Edition). Boca Raton: CRC Press, Taylor & Francis Group.

O'BRIEN, E.; ZHANG, X.; NISHIMURA, C.; TOMBLIN, J. B. & MURRAY, J. (2003). Association of specific language impairment (SLI) to the region of 7q31. *American Journal of Human Genetics*, 72, pp. 1536-1543.

ORSINI, J. J.; ESCOLAR, M. L.; WASSERSTEIN, M. P. *et al.* (2000 [Updated 2018]). Krabbe Disease. In: ADAM, M. P.; ARDINGER, H. H.; PAGON, R. A. *et al.* (eds.). *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle.

PATTERSON, F. G. (1978). The gestures of a gorilla: language acquisition in another pongid. *Brain and language*, 5(1), pp. 72-97.

PEPPERBERG, I. M. (1990). Cognition in an African Grey parrot (*Psittacus erithacus*): Further evidence for comprehension of categories and labels. *Journal of Comparative Psychology*, 104, pp. 41-52.

_____. (1992). Proficient performance of a conjunctive, recursive task by an African Grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, 106, pp. 295-305.

_____. (1999). *The Alex studies: cognitive and communicative abilities of grey parrots*. Cambridge, Massachusetts; London, England: Harvard University Press.

PETRYSHEN, T. L.; KAPLAN, B. J.; FU LIU, M.; DE FRENCH, N. S.; TOBIAS, R.; HUGHES, M. L. & FIELD, L. L. (2001). Evidence for a susceptibility locus on chromosome 6q influencing phonological coding dyslexia. *American journal of medical genetics*, 105(6), pp. 507-517.

PFEIFER, L. S.; SCHMITZ, J.; PAPADATOU-PASTOU, M.; PETERBURS, J.; PARACCHINI, S. & OCKLENBURG, S. (2022). Handedness in twins: meta-analyses. *BMC Psychology* 10, 11, pp. 1-19.

PFENNING, A. R.; HARA, E.; WHITNEY, O.; RIVAS, M. V.; WANG, R.; ROULHAC, P. L.; HOWARD, J. T.; WIRTHLIN, M.; LOVELL, P. V.; GANAPATHY, G.; MOUNCASTLE, J.; MOSELEY, M. A.; THOMPSON, J. W.; SODERBLOM, E. J.; IRIKI, A.; KATO, M.; GILBERT, M. T.; ZHANG, G.; BAKKEN, T.; BONGAARTS, A. & JARVIS, E. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* (New York, N.Y.), 346(6215), 1256846.

PHELAN, M. C.; ROGERS, R. C.; SAUL, R. A.; STAPLETON, G. A.; SWEET, K.; MCDERMID, H.; SHAW, S. R.; CLAYTOR, J.; WILLIS, J. & KELLY, D. P. (2001). 22q13 deletion syndrome. *American journal of medical genetics*, 101(2), pp. 91-99.

PIAGET, J. (1975). *Epistemologia genética*. São Paulo: Abril Cultural.

_____. (1979). *A construção do real na criança*. Rio de Janeiro: Zahar Editores.

PIATELLI-PALMARINI, M. (ed.) (1980). *Language and learning: the debate between Jean Piaget and Noam Chomsky*. Cambridge: The MIT Press.

PIATELLI-PALMARINI, M. & URIAGEREKA, J. (2011). A Geneticist's Dream, a Linguist's Nightmare: The Case of FOXP2. In: DI SCIULLO, A. M. & BOECKX, C. (eds.).

The biolinguistic enterprise: new perspectives on the evolution and nature of the human language faculty. Oxford: Oxford University Press.

PIERCE, B. A. (2016). *Genética: um enfoque conceitual* (5ª edição). Tradução de Beatriz Araujo do Rosário. Rio de Janeiro: Guanabara Koogan.

PINKER, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

PINKER, S. (1994). *The language instinct.* New York: Morrow.

PLANTE, E.; SHENKMAN, K. & CLARK, M. (1996). Classification of adults for family studies of developmental language disorders. *Journal of Speech and Hearing Research*, 39, pp. 661-667.

PLOMIN, R. & RUTTER, M. (1998). Child development, molecular genetics, and what to do with genes once they are found. *Child development*, 69(4), pp. 1221-1240.

POEPPPEL, D. (2001). Pureword deafness and the bilateral processing of the speech code. *Cogn. Sci.* 25, pp. 679-693.

POEPPPEL, D. & EMBICK, D. (2006). The Relation Between Linguistics and Neuroscience. In: CUTLER, A. (ed.). *Twenty-first Century Psycholinguistics: Four Cornerstones.* Mahwah: Lawrence Erlbaum.

PORTER, F. D. (2008). Smith-Lemli-Opitz syndrome: pathogenesis, diagnosis and management. *European journal of human genetics*, 16(5), pp. 535-541.

PORTER, J. A.; YOUNG, K. E. & BEACHY, P. A. (1996). Cholesterol modification of hedgehog signaling proteins in animal development. *Science*, 274(5285), pp. 255-259.

PRASAD, C.; PRASAD, A. N.; CHODIRKER, B. N.; LEE, C.; DAWSON, A. K.; JOCELYN, L. J. & CHUDLEY, A. E. (2000). Genetic evaluation of pervasive developmental disorders: the terminal 22q13 deletion syndrome may represent a recognizable phenotype. *Clinical genetics*, 57(2), pp. 103-109.

PROCACCIO, V. & WALLACE, D. C. (2004). Late-onset Leigh syndrome in a patient with mitochondrial complex I NDUF8 mutations. *Neurology*, 62(10), pp. 1899-1901.

RABIN, M.; WEN, X. L.; HEPBURN, M.; LUBS, H. A.; FELDMAN, E. & DUARA, R. (1993). Suggestive linkage of developmental dyslexia to chromosome 1p34-p36. *Lancet*, 342, p. 178.

RAHMAN, S. & THORBURN, D. (2015 [Updated 2020]). Nuclear Gene-Encoded Leigh Syndrome Spectrum Overview. In: ADAM, M.P.; ARDINGER, H. H.; PAGON, R. A. *et al.* (eds). *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle.

RAPIN, I. D. & ALLEN, D. A. (1988). Syndromes in developmental dysphasia and adultaphasia. In: PLUM, F. (ed.). *Language. Communication and the brain*. New York: Raven Press.

RAPOSO, E. P. (1992). *Teoria da gramática: a faculdade da linguagem*. Lisboa: Caminho.

REFETOFF, S.; WEISS, R. E. & USALA, S. J. (1993). The syndromes of resistance to thyroid hormone. *Endocrine reviews*, 14(3), pp. 348–399.

RICE, M. (2012). Toward epigenetic and gene regulation models of specific language impairment: Looking for links among growth, genes, and impairments. *Journal of Neurodevelopmental Disorders*, 4(27), pp. 1-14.

RICE, M.; HANEY, K. & WEXLER, K. (1998). Family histories of children with SLI who show extended optional infinitives. *Journal of Speech, Language and Hearing Research*, 41, pp. 419-432.

RICE, M.; SMITH, S. & GAYÁN, J. (2009). Convergent genetic linkage and associations to language, speech, and reading measures in families of probands with Specific Language Impairment. *Journal of Neurodevelopmental Disorders*, 1, pp. 264-282.

RICE, M. . & WEXLER, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of speech and hearing research*, 39(6). Res. 39, pp. 1239-1257.

RIZZI, L. (1989). The new comparative syntax: principles and parameters of Universal Grammar. *Cahiers Ferdinand De Saussure*, n. 43, pp. 65-78.

ROEBROEKS, W. & SORESSI, M. (2016). Neandertals revised. *Proceedings of the National Academy of Sciences of the United States of America*, 113(23), pp. 6372-6379.

ROLL, P.; RUDOLF, G.; PEREIRA, S.; ROYER, B.; SCHEFFER, I. E.; MASSACRIER, A.; VALENTI, M. P.; ROECKEL-TREVISIOL, N.; JAMALI, S.; BECLIN, C.; SEEGMULLER, C.; METZ-LUTZ, M. N.; LEMAINQUE, A.; DELEPINE, M.; CALOUSTIAN, C.; DE SAINT MARTIN, A.; BRUNEAU, N.; DEPETRIS, D.; MATTEI, M. G.; FLORI, E.; ROBAGLIA-SCHLUPP, A.; LEVY, N.; NEUBAUER, B. A.; RAVID, R.; MARESCAUX, C.; BERKOVIC, S. F.; HIRSCH, E.; LATHROP, M.; CAU, P. & SZEPETOWSKI, P. (2006). SRPX2 mutations in disorders of language cortex and cognition. *Human molecular genetics*, 15(7), pp. 1195-1207.

ROMAN, T.; SCHMITZ, M.; POLANCZYK, G.; EIZIRIK, M.; ROHDE, L. A. & HUTZ, M. H. (2001). Attention-deficit hyperactivity disorder: a study of association with both the dopamine transporter gene and the dopamine D4 receptor gene. *American journal of medical genetics*, 105(5), pp. 471-478.

ROSS, J. R. (1967). *Constraints on variables in Syntax*. Tese de doutorado. MIT.

RUBINSZTEIN, D. C.; LEGGO, J.; CHIANO, M.; DODGE, A.; NORBURY, G.; ROSSER, E. & CRAUFURD, D. (1997). Genotypes at the GluR6 kainate receptor locus are associated with variation in the age of onset of Huntington disease. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8), pp. 3872-3876.

SALOMONS, G. S.; van DOOREN, S. J.; VERHOEVEN, N. M.; MARSDEN, D.; SCHWARTZ, C.; CECIL, K. M.; DeGRAUW, T. J. & JAKOBS, C. (2003). X-linked creatine transporter defect: an overview. *Journal of inherited metabolic disease*, 26(2-3), pp. 309-318.

SCALLY, A.; DUTHEIL, J. Y.; HILLIER, L. W.; JORDAN, G. E.; GOODHEAD, I.; HERRERO, J.; HOBOLTH, A.; LAPPALAINEN, T.; MAILUND, T.; MARQUES-BONET, T.; MCCARTHY, S.; MONTGOMERY, S. H.; SCHWALIE, P. C.; TANG, Y. A.; WARD, M. C.; XUE, Y.; YNGVADOTTIR, B.; ALKAN, C.; ANDERSEN, L. N.; AYUB, Q. & DURBIN, R. (2012). Insights into hominid evolution from the gorilla genome sequence. *Nature*, 483(7388), pp. 169-175.

SCHARFF, C. & HAESLER, S. (2005). An evolutionary perspective on FoxP2: strictly for the birds?. *Current opinion in neurobiology*, 15(6), pp. 694-703.

SCHEFFER, I. E.; JONES, L.; POZZEBON, M.; HOWELL, R. A.; SALING, M. M. & BERKOVIC, S. F. (1995). Autosomal dominant rolandic epilepsy and speech dyspraxia: a new syndrome with anticipation. *Annals of neurology*, 38(4), pp. 633-642.

SCHUMACHER, J.; ANTHONI, H.; DAHDOUH, F.; KÖNIG, I. R.; HILLMER, A. M.; KLUCK, N.; MANTHEY, M.; PLUME, E.; WARNKE, A.; REMSCHMIDT, H.; HÜLSMANN, J.; CICHON, S.; LINDGREN, C. M.; PROPPING, P.; ZUCHELLI, M.; ZIEGLER, A.; PEYRARD-JANVID, M.; SCHULTE-KÖRNE, G.; NÖTHEN, M. M. & KERE, J. (2006). Strong genetic evidence of DCDC2 as a susceptibility gene for dyslexia. *American journal of human genetics*, 78(1), pp. 52-62.

SCOTT, H. S.; BLANCH, L.; GUO, X. H.; FREEMAN, C.; ORSBORN, A.; BAKER, E.; SUTHERLAND, G. R.; MORRIS, C. P. & HOPWOOD, J. J. (1995). Cloning of the sulphamidase gene and identification of mutations in Sanfilippo A syndrome. *Nature genetics*, 11(4), pp. 465-467.

SCOTT, H. S.; BUNGE, S.; GAL, A.; CLARKE, L. A.; MORRIS, C. P. & HOPWOOD, J. J. (1995). Molecular genetics of mucopolysaccharidosis type I: diagnostic, clinical, and biological implications. *Human mutation*, 6(4), pp. 288-302.

SEDEL, F.; FRIDERICI, K.; NUMMY, K.; CAILLAUD, C.; CHABLI, A.; DÜRR, A.; LUBETZKI, C. & AGID, Y. (2006). Atypical Gilles de la Tourette Syndrome with beta-mannosidase deficiency. *Archives of neurology*, 63(1), pp. 129-131.

SHAPIRO, E. G.; WHITLEY, C. B. & EISENGART, J. B. (2018). Beneath the floor: re-analysis of neurodevelopmental outcomes in untreated Hurler syndrome. *Orphanet Journal of Rare Diseases*, 13, p. 76.

SHIH, L. Y.; SUSLAK, L.; ROSIN, I.; SEARLE, B. M. & DESPOSITO, F. (1984). Gene dosage studies supporting localization of the structural gene for galactose-1-phosphate uridyl transferase (GALT) to band p13 of chromosome 9. *American journal of medical genetics*, 19(3), pp. 539-543.

SHRIBERG, L. D.; BALLARD, K. J.; TOMBLIN, J. B.; DUFFY, J. R.; ODELL, K. H. & WILLIAMS, C. A. (2006). Speech, prosody, and voice characteristics of a mother and daughter with a 7,13 translocation affecting FOXP2. *Journal of speech and hearing research*. 49, pp. 500-525.

SHU, W.; YANG, H.; ZHANG, L.; LU, M. M. & MORRISEY E. E. (2001). Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. *The Journal of biological chemistry*, 276 (29), pp. 27488–27497.

SIKANSI, N. S. (1999). As interrogativas-Q na gramática infantil do PB. *Caderno de Estudos Linguísticos*, Campinas, SP, v.36, Jan/Jun. 1999, pp. 85-103.

SLI Consortium (SLIC) (2002). A genomewide scan identifies two novel loci involved in specific language impairment. *American journal of human genetics*, 70, pp. 384-398.

SLI Consortium (SLIC) (2004). Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *American journal of human genetics*, 74, pp. 1225-1238.

SNYDER, M. & GERSTEIN, M. (2003). Defining genes in the genomics era. *Science*, 300, pp. 258-260.

SONNINEN, P.; AUTTI, T.; VARHO, T., HÄMÄLÄINEN, M. & RAININKO, R. (1999). Brain involvement in Salla disease. *American journal of neuroradiology*, 20(3), pp. 433-443.

SPLAWSKI, I.; TIMOTHY, K. W.; SHARPE, L. M.; DECHER, N.; KUMAR, P.; BLOISE, R.; NAPOLITANO, C.; SCHWARTZ, P. J.; JOSEPH, R. M.; CONDOURIS, K.; TAGER-FLUSBERG, H.; PRIORI, S. G.; SANGUINETTI, M. C. & KEATING, M. T. (2004). Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell*, 119, pp. 19-31.

STAES, N.; SHERWOOD, C. C.; WRIGHT, K.; DE MANUEL, M.; GUEVARA, E. E.; MARQUES-BONET, T.; KRÜTZEN, M.; MASSIAH, M.; HOPKINS, W. D.; ELY, J. J. & BRADLEY, B. J. (2017). FOXP2 variation in great ape populations offers insight into the evolution of communication skills. *Scientific reports*, 7(1), p. 16866.

STEIN, C. M.; MILLARD, C.; KLUGE, A.; MISCIMARRA, L. E.; CARTIER, K. C.; FREEBAIRN, L. A.; HANSEN, A. J.; SHRIBERG, L. D.; TAYLOR, H. G.; LEWIS, B. A. & IYENGAR, S. K. (2006). Speech sound disorder influenced by a locus in 15q14 region. *Behavior genetics*, 36(6), pp. 858-868.

STEIN, C. M.; SCHICK, J. H.; TAYLOR, H. G.; SHRIBERG, L. D.; MILLARD, C.; KUNDTZ-KLUGE, A.; RUSSO, K.; MINICH, N.; HANSEN, A.; FREEBAIRN, L. A.; ELSTON, R. C.; LEWIS, B. A. & IYENGAR, S. K. (2004). Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. *American journal of human genetics*, 74(2), pp. 283–297.

SUN, T.; PATOINE, C.; ABU-KHALIL, A.; VISVADER, J.; SUM, E.; CHERRY, T. J.; ORKIN, S. H.; GESCHWIND, D. H. & WALSH, C. A. (2005). Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science*, 308(5729), pp. 1794–1798.

SWANBERG, M.; NASREDDINE, Z.; MENDEZ, M. & CUMMINGS, J. (2007). Speech and Language. In: GOETZ, C. G. (ed.). *Textbook of Clinical Neurology* (Third Edition). Philadelphia: W.B. Saunders, pp. 79–98.

TAIPALE, M.; KAMINEN, N.; NOPOLA-HEMMI, J.; HALTIA, T.; MYLLYLUOMA, B.; LYYTINEN, H.; MULLER, K.; KAARANEN, M.; LINDSBERG, P. J.; HANNULA-JOUPPI, K. & KERE, J. (2003). A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. *Proceedings of the National Academy of Sciences of the United States of America*, 100(20), pp. 11553–11558.

TALLAL, P.; HIRSCH, L.; REALPE-BONILLA, T.; MILLER, S.; BRZUSTOWICZ, L.; BARTLETT, C. et al. (2001). Familial aggregation in specific language impairment. *Journal of Speech, Language and Hearing Research*, 44, pp. 1172–1182.

TALLAL, P.; ROSS, R. & CURTISS, S. (1989). Familial aggregation in specific language impairment. *Journal of Speech and Hearing Disorders*, 54, pp. 167–173.

TALLAL, P.; SAINBURG, R. L. & JERNIGAN, T. (1991). The Neuropathology of Developmental Dysphasia: Behavioral, Morphological, and Physiological Evidence for a Pervasive Temporal Processing Disorder. *Reading and writing*, 3(3–4), pp. 363–377.

TEICHMANN, M.; DUPOUX, E.; KOUIDER, S.; BRUGIERES, P.; BOISSE, M. F.; BAUDIC, S.; CESARO, P.; PESCHANSKI, M. & BACHOUD-LEVI, A. C. (2005). The role of the striatum in rule application: the model of Huntington's disease at early stage. *Brain*, 128, pp. 1155–1167.

TERAMITSU, I.; KUDO, L. C.; LONDON, S. E.; GESCHWIND, D. H. & WHITE, S. A. (2004). Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 24(13), pp. 3152-3163.

TERRACE, H. S., PETITTO, L. A., SANDERS, R. J. & BEVER, T. G. (1979). Can an ape create a sentence? *Science*, 206 (4421), pp. 891-902.

THORPE, J. E. (1988). Salmon migration. *Science Progress*, v.72, n.3 (287), pp. 345-370.

TIAN, M.; QU, Y.; HUANG, L. *et al.* (2019). Stable clinical course in three siblings with late-onset isolated sulfite oxidase deficiency: a case series and literature review. *BMC Pediatrics*, 19, p. 510.

TIERNEY, E.; NWOKORO, N. A.; PORTER, F. D.; FREUND, L. S.; GHUMAN, J. K. & KELLEY, R. I. (2001). Behavior phenotype in the RSH/Smith-Lemli-Opitz syndrome. *American journal of medical genetics*, 98(2), pp. 191-200.

TIMMERS, I.; van den HURK, J.; DI SALLE, F.; RUBIO-GOZALBO, M. E. & JANSMA, B. M. (2011). Language production and working memory in classic galactosemia from a cognitive neuroscience perspective: future research directions. *Journal of inherited metabolic disease*, 34(2), pp. 367-376.

TIMSON, D. J. (2016). The molecular basis of galactosemia: past, present and future. *Gene*, 589 (2), pp. 133-141.

TOMBLIN, J. B. (1989). Familial concentration of developmental language impairment. *Journal of Speech and Hearing Disorders*, 54, pp. 287-295.

TOMBLIN, J. B. & BUCKWALTER, P. (1994). Studies of genetics of specific language impairment. In: WATKINS, R. V. & RICE, M. L. (Eds.). *Specific language impairments in children*. Baltimore: Paul H. Brookes.

TOMBLIN, J. B. & BUCKWALTER, P. (1998). Heritability of poor language achievement among twins. *Journal of speech, language and hearing research*, 41, pp. 188-199.

TURNBAUGH, P. J.; LEY, R. E.; HAMADY, M.; FRASER-LIGGETT, C. M.; KNIGHT, R. & GORDON, J. I. (2007). The Human Microbiome Project. *Nature*, 449 (7164), pp. 804-810.

UCHINO, S. & WAGA, C. (2013). SHANK3 as an autism spectrum disorder-associated gene. *Brain & development*, 35(2), pp. 106-110.

UNIVERSIDADE DE SÃO PAULO (USP). Instituto de Biociências da USP. *Biologia Molecular: O código genético*.

VALSTAR, M. J.; NEIJS, S.; BRUGGENWIRTH, H. T.; OLMER, R.; RUIJTER, G. J.; WEVERS, R.A.; VAN DIGGELEN, O. P.; POORTHUIS, B. J.; HALLEY, D. J. & WIJBURG, F. A. (2010). Mucopolysaccharidosis type IIIA: clinical spectrum and genotype-phenotype correlations. *Annals of neurology*, 68(6), pp. 876-887.

VAN DE KAMP, J. J.; NIERMEIJER, M. F.; VON FIGURA, K. & GIESBERTS, M. A. (1981). Genetic heterogeneity and clinical variability in the Sanfilippo syndrome (types A, B, and C). *Clinical genetics*, 20(2), pp. 152-160.

VAN DEN BROEK, B.; SMIT, A. L.; BOELEN, J. J. & van HASSELT, P. M. (2020). Hearing loss in patients with mucopolysaccharidoses-1 and 6 after hematopoietic cell transplantation: A longitudinal analysis. *Journal of inherited metabolic disease*, 43(6), pp. 1279-1287.

VAN DER LELY, H. & STOLLWERCK, L. (1996). A grammatical specific language impairment in children: An autosomal dominant inheritance?. *Brain and Language*, 52, pp. 484-504.

VAN DIGGELEN, O. P.; SCHINDLER, D.; KLEIJER, W. J.; HUIJMANS, J. M.; GALJAARD, H.; LINDEN, H. U.; PETER-KATALINIC, J.; EGGE, H.; DABROWSKI, U. & CANTZ, M. (1987). Lysosomal alpha-N-acetylgalactosaminidase deficiency: a new inherited metabolic disease. *Lancet*, 2(8562), p. 804.

VAN HOVE, J. L.; COUGHLIN, C.; SWANSON, M. *et al.* (2002 [Updated 2019]). Nonketotic Hyperglycinemia. In: ADAM, M.P.; ARDINGER, H. H.; PAGON, R. A. *et al.* (eds). *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle.

VAN RAPPARD, D. F.; BOELEN, J. J. & WOLF, N. I. (2015). Metachromatic leukodystrophy: Disease spectrum and approaches for treatment. Best practice & research. *Clinical endocrinology & metabolism*, 29(2), pp. 261-273.

VARGHA-KHADEM, F.; GADIAN, D. G.; COPP, A. & MISHKIN, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature reviews. Neuroscience*, 6(2), pp. 131-138.

VARGHA-KHADEM, F. & PASSINGHAM, R. E. (1990). Speech and language defects. *Nature*, 346, p. 226.

VARGHA-KHADEM, F.; WATKINS, K.; ALCOCK, K.; FLETCHER, P.; PASSINGHAM, R. E. (1995). Cognitive and praxic deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, 92(3), pp. 930-933.

VARGHA-KHADEM, F.; WATKINS, K. E.; PRICE, C. J.; ASHBURNER, J.; ALCOCK, K. J.; CONNELLY, A.; FRACKOWIAK, R. S.; FRISTON, K. J.; PEMBREY, M. E.; MISHKIN, M.; GADIAN, D. G. & PASSINGHAM, R. E. (1998). Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, 95(21), pp. 12695-12700.

VARHO, T. T.; ALAJOKI, L. E.; POSTI, K. M.; KORHONEN, T. T.; RENLUND, M. G. NYMAN, S. R. G.; SILLANPÄÄ M. L. & AULA, P. P. (2002). Phenotypic spectrum of Salla disease, a free sialic acid storage disorder. *Pediatric Neurology*, v.26, n.4, pp. 267-273.

VENTER, J. C.; ADAMS, M. D.; MYERS, E. W.; LI, P. W.; MURAL, R. J.; SUTTON, G. G.; SMITH, H. O.; YANDELL, M.; EVANS, C. A.; HOLT, R. A.; GOCAYNE, J. D.; AMANATIDES, P.; BALLEW, R. M.; HUSON, D. H.; WORTMAN, J. R.; ZHANG, Q.; KODIRA, C. D.; ZHENG, X. H.; CHEN, L.; SKUPSKI, M.; ... ZHU, X. (2001). The sequence of the human genome. *Science*, 291(5507), pp. 1304-1351.

VERNES, S. C.; NICOD, J.; ELAHI, F. M.; COVENTRY, J. A.; KENNY, N.; COUPE, A. M.; BIRD, L. E.; DAVIES, K. E. & FISHER, S. E. (2006). Functional genetic analysis of mutations implicated in a human speech and language disorder. *Human molecular genetics*, 15(21), pp. 3154–3167.

VOIT, T.; BELTRAN-VALERO de BERNABÉ, D.; LONGMAN, C.; STEINBRECHER, A.; STRAUB, V.; YUVA, Y.; HERRMANN, R.; SPERNER, J.; KORENKE, C.; DIESEN, C.; DOBYNS, W. B.; BRUNNER, H. G.; van BOKHOVEN, H.; BROCKINGTON, M. & MUNTONI, F. (2004). Mutations in the FKRP gene can cause muscle-eye-brain disease and Walker-Warburg syndrome. *Journal of medical genetics*, 41(5): e61.

WALKER, R. H.; MORGELLO, S.; DAVIDOFF-FELDMAN, B.; MELNICK, A.; WALSH, M. J.; SHASHIDHARAN, P. & BRIN, M. F. (2002). Autosomal dominant chorea-acanthocytosis with polyglutamine-containing neuronal inclusions. *Neurology*, 58, pp. 1031-1037.

WANG, D.; PASCUAL, J. M.; YANG, H.; ENGELSTAD, K.; JHUNG, S.; SUN, R. P. & DE VIVO, D. C. (2005). Glut-1 deficiency syndrome: clinical, genetic, and therapeutic aspects. *Annals of neurology*, 57(1), pp. 111-118.

WANG, H.; WANG, A.; WANG, D.; BRIGHT, A.; SENCY, V.; ZHOU, A. & XIN, B. (2016). Early growth and development impairments in patients with ganglioside GM3 synthase deficiency. *Clinical genetics*, 89(5), pp. 625-629.

WANG, Y.; KORMAN, S. H.; YE, J.; GARGUS, J. J.; GUTMAN, A.; TARONI, F.; GARAVAGLIA, B. & LONGO, N. (2001). Phenotype and genotype variation in primary carnitine deficiency. *Genetics in medicine: official journal of the American College of Medical Genetics*, 3(6), pp. 387-392.

WATKINS, K. E.; DRONKERS, N. F. & VARGHA-KHADEM, F. (2002) Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*, 125, pp. 452-464.

WEBB, A. L.; SINGH, R. H.; KENNEDY, M. J. & ELSAS, L. J. (2003). Verbal dyspraxia and galactosemia. *Pediatric Research*, 53, pp. 396-402.

WENGER, D. A.; RAFI, M. A. & LUZI, P. (2016). Krabbe disease: One Hundred years from the bedside to the bench to the bedside. *Journal of neuroscience research*, 94(11), pp. 982-989.

WHITE, D. A.; NORTZ, M. J.; MANDERNACH, T.; HUNTINGTON, K. & STEINER, R. D. (2001). Deficits in memory strategy use related to prefrontal dysfunction during early development: evidence from children with phenylketonuria. *Neuropsychology*, 15, pp. 221-229.

WILSON, H. L.; WONG, A. C.; SHAW, S. R.; TSE, W. Y.; STAPLETON, G. A.; PHELAN, M. C.; HU, S.; MARSHALL, J. & McDERMID, H. E. (2003). Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. *Journal of medical genetics*, 40(8), pp. 575-584.

WU, W.; MORENO, A. M.; TANGEN, J. M. & REINHARD, J. (2013). Honeybees can discriminate between Monet and Picasso paintings. *Journal of comparative physiology A. Neuroethology, sensory, neural, and behavioral physiology*, 199(1), pp. 45-55.

XIANG, F.; ALMQVIST, E. W.; HUQ, M.; LUNDIN, A.; HAYDEN, M. R.; EDSTRÖM, L.; ANVRET, M. & ZHANG, Z. (1998). A Huntington disease-like neurodegenerative disorder maps to chromosome 20p. *American journal of human genetics*, 63(5), pp. 1431-1438.

YOGALINGAM, G. & HOPWOOD, J. J. (2001). Molecular genetics of mucopolysaccharidosis type IIIA and IIIB: Diagnostic, clinical, and biological implications. *Human mutation*, 18(4), pp. 264-281.