

ALÉXIA TELES DUCHOWNY *ET AL.*

APOSTILA:
FUNDAMENTOS DE LINGUÍSTICA COMPARADA
PRESENCIAL

Texto 1: O que é linguística comparada? p. 2
 Guia de leitura p. 11
Texto 2: Arqueologias p. 12
 Guia de leitura p. 21
Texto 3: O método histórico-comparativo p. 22
 Guia de leitura p. 39
Texto 4: A reconstrução do indo-europeu p. 40
 Guia de leitura p. 65
Texto 5: O que é uma língua p. 66
 Guia de leitura p. 92
Texto 6: As línguas do mundo p. 93
 Guia de leitura p. 113
Texto 7: Sistemas de escrita p. 115
 Guia de leitura p. 124
Texto 8: As línguas indo-europeias p. 125
 Guia de leitura p. 140
Texto 9: As línguas da África p. 142
 Guia de leitura p. 163
Texto 10: As línguas indígenas brasileiras p. 164
 Guia de leitura p. 182
Bibliografia p. 183

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Faculdade de Letras

Belo Horizonte

janeiro/2015

TEXTO 1

O QUE É LINGUÍSTICA COMPARADA?

Jacyntho Lins Brandão

Houve uma vez que os estudantes de Letras da UFMG mandaram confeccionar camisas com os dizeres:

THE BOOK IS ON THE TABLE
LE LIVRE EST SUR LA TABLE
EL LIBRO ESTÁ SOBRE LA MESA
DAS BUCH IST AUF DEM TISCH
ΤΟ ΒΙΒΛΙΟΝ ΕΣΤΙΝ ΕΠΙ ΤΗ ΤΡΑΠΕΖΗ
IL LIBRO È SULLA TAVOLA
LIBER SUPER MENSAM EST
O LIVRO ESTÁ SOBRE A MESA

Como você vê, uma brincadeira divertida com uma das frases que mais costumava aparecer em livros tradicionais para ensino de línguas estrangeiras, sobretudo o inglês. As oito línguas faziam parte do rol das habilitações ofertadas na Faculdade de Letras, envolvendo, além das modernas, duas antigas (o latim e o grego clássico) e dois alfabetos diferentes (o grego e o latino). No caso do grego, a transliteração para o alfabeto latino é:

Τὸ βιβλίον ἐστὶν ἐπὶ τῆι τραπέζῃ.¹

Mas mesmo com essa diversidade, não era difícil entender que as frases correspondiam umas às outras praticamente palavra por palavra. Antes de tudo, porque todas são línguas de uma mesma família, a indo-europeia, representada na relação por três de suas dez ramificações: o itálico (com o latim e as quatro línguas dele procedentes: francês, espanhol,

¹ A duração das vogais, quando for importante marcá-la, será indicada assim: (a) vogais longas: ā/ē/ī/ō/ū (o traço horizontal sobre elas se chama *macro*); (b) vogais breves: ă/ĕ/ĭ/ŏ/ŭ (o símbolo sobre elas se chama *braquiá*).

italiano e português, chamadas de línguas românicas); o germânico (com o inglês e o alemão); e o grego.

Na simples comparação, considerando a ordem das palavras e sua semelhança, não será difícil que você descubra a correspondência entre as mesmas. Experimente:

QUADRO 1
O livro está sobre a mesa

Português	<i>o</i>	<i>livro</i>	<i>está</i>	<i>sobre</i>	<i>a</i>	<i>mesa</i>
Inglês						
Francês						
Espanhol						
Alemão						
Grego						
Italiano						
Latim						

Você deve ter encontrado dois problemas:

1. Com relação ao italiano, *sulla* constitui uma contração da preposição *su* e do artigo *la* (do mesmo modo que, em português, temos *da* < *de a* e *na* < *em a*).²
2. Você deve ter notado que o latim não possui artigos e adota uma ordem diferente dos termos da oração: em vez de

sujeito verbo locativo
o livro está sobre a mesa,

a ordem normal em latim é

sujeito locativo verbo
liber super mensam est.

Tendo constatado essas duas diferenças, apenas para que a correspondência no quadro se faça palavra a palavra, anote a preposição *su* separada do artigo *la*, no caso do italiano, e

² Aos poucos você se acostumará com os símbolos que utilizaremos: B < A indica que a palavra B procede de A, o que pode ser representado também assim, A > B. A ordem da procedência segue sempre a direção indicada pela seta.

escreva os termos latinos na ordem dos demais, deixando em branco os espaços em que as outras línguas apresentam artigos.

Agora observe na segunda coluna as palavras que designam 'livro'. Não será difícil constatar que elas se distribuem em três conjuntos, considerando-se sua semelhança: o primeiro, com cinco termos; o segundo, com dois; e o terceiro, com apenas um. Organize-os:

QUADRO 2
Palavras para 'livro'

1	<i>livro</i>
2	
3	

Repare que essa distribuição corresponde exatamente às famílias linguísticas referidas antes: 1. latim e as línguas românicas; 2. línguas germânicas; 3. grego. As semelhanças, portanto, não são fortuitas, mas decorrem do fato de que:

- (a) O português, o francês, o espanhol e o italiano originaram-se do latim;
- (b) O inglês e o alemão têm uma origem comum;
- (c) O grego constitui um grupo isolado dentre as demais línguas indo-europeias.

Mesmo que nos três grupos as palavras para designar 'livro' sejam diferentes, têm elas em comum o fato de que, na origem, nomeavam o material sobre o qual se escrevia:

1. O termo grego *biblíon* deriva de *byblos*, 'papiro', a planta natural do Egito com a qual se produzia a folha (em grego *khárta*) em que se escrevia e com as quais se produziam os livros. O plural *biblíia* passou para as línguas modernas como nome do conjunto de livros sagrados de judeus e cristãos, a Bíblia.
2. Para *book* e *Buch*³ reconstituiu-se, no germânico, a palavra **bōks*⁴, relacionada com **bōka*, 'faia', porque os povos germânicos usavam cascas dessa árvore para escrever. De **bōks* provêm os termos do inglês antigo *bōc* (donde, por sua vez, procede *book*), alemão *Buch*, holandês *boek*, sueco *bok* etc, todos significando 'livro'.
3. A palavra latina *liber* significa originalmente 'casca', a 'entrecasca' em que se escrevia antes da adoção do papiro, passando a nomear, em seguida, o próprio livro. Os termos das línguas românicas procedem do acusativo de *liber*, ou seja,

³ Em alemão, todos os substantivos se escrevem com inicial maiúscula: *Buch*, *Tisch* etc.

⁴ As palavras marcadas com um asterisco não são documentadas, mas reconstituídas, pelo método comparativo. Isso se faz sistematicamente nos estudos de linguística histórico-comparativa.

librum, o qual dá origem a *libro*, em italiano e espanhol, *livro*, em português, e *livre*, em francês.⁵

Examine agora a última coluna do nosso quadro, agrupando os termos de acordo com sua semelhança. Você encontrará quatro tipos, numa distribuição diferente da anterior:

QUADRO 3

Palavras para 'mesa'

<i>mensam</i>	<i>table</i>	<i>Tisch</i>	<i>trapézēi</i>
---------------	--------------	--------------	-----------------

Fica claro que as palavras do espanhol e do português, *mesa*, provêm do latim *mensa*.

Entretanto, também as palavras da segunda coluna têm uma origem latina: *tabŭla* significa 'tábua', 'tabuleiro', estando na origem do italiano *tavola* e do francês *table* (o português tem, da mesma origem, o termo *tábua*, assim como o espanhol, *tabla*; recorde-se ainda que *távola*, com o significado de 'mesa', existe também em português, embora seja um arcaísmo, fossilizado, por exemplo, na referência ao Rei Artur e "os cavaleiros da tábua redonda").

O inglês *table* procede do francês, por empréstimo, como acontece com grande parte do vocabulário daquela língua, em consequência do domínio normando, iniciado em 1066, sobre as Ilhas Britânicas.

O caso do alemão também se deve a um longo processo de empréstimos: (a) o termo original é o grego *dískos*, 'disco, objeto circular', 'disco de arremesso' (como continua a ser usado nos jogos olímpicos); (b) o latim *discus*, 'prato', 'travessa redonda', constitui um empréstimo da palavra grega citada; (c) o germânico tomou emprestado o termo latino, **disku/diskuz*, significando 'prato', 'travessa', 'tábua de comida', 'bandeja', 'mesa', donde provém a palavra do antigo-alto-alemão *tisk/tisc*, 'mesa', 'prato', 'travessa', 'trípode', 'bandeja', origem do termo do médio-alto-alemão *tisch*, 'mesa', 'mesa onde se come' e do alemão *Tisch*, 'mesa'. É curioso que em alemão existe também uma outra palavra para 'mesa', 'tábua': *Tafel*, que procede do médio-alto-alemão *tavele/tabele*, por sua vez proveniente do antigo-alto-alemão *tavala/tabala*, empréstimo do latim *tabŭla*. Veja como esses fatos linguísticos sugerem que os germanos não possuíam uma palavra para 'mesa' e parecem ter tomado dos romanos tanto o objeto, quanto sua denominação.

⁵ Acusativo é a forma que a palavra assume quando se encontra na função de objeto ou regida por certas preposições. A forma da palavra quando está na função de sujeito se chama nominativo: *liber* (nominativo)/*librum* (acusativo); *mensa* (nominativo)/*mensam* (acusativo).

Enfim, o grego *trápeza* é uma abreviação de *tetrapéza*, isto é, 'de quatro pés', o que remete para a forma da mesa.

* * *

Não continuaremos a explorar a origem dos demais termos de nossas oito frases, aos quais teremos oportunidade de voltar nas lições seguintes. Uma vez que lidamos apenas com línguas indo-europeias, tudo que fizemos até aqui teve uma perspectiva histórica. Esse será um dos vetores de nosso curso, o qual abordaremos na primeira metade, ao tratarmos dos primórdios da linguística comparada – que se deu com o estudo da família indo-europeia – e das diferentes famílias linguísticas do mundo.

Outro método de comparação das línguas, que gera uma classificação diferente da genética, é o da linguística tipológica ou tipologia linguística. Este será o assunto da segunda metade do nosso curso, quando não mais nos interessarão as relações entre línguas de uma mesma família, mas os traços gramaticais que sejam comuns a certas línguas.

Para adiantar essa forma de tratamento, continuando a utilizar a nossa frase padrão, nas suas oito versões, observaremos agora três aspectos: (a) o uso de artigos; (b) a flexão nominal; (c) o uso de preposições.

Você já observou que o latim não possui artigo definido, tanto que deixou em branco os espaços do quadro 1 ocupados pelos artigos das demais línguas. Agora observe que esses artigos, nas línguas que os possuem, têm uma ou mais formas, como se mostra no quadro seguinte:

QUADRO 4

Línguas sem artigo/línguas com artigo definido

Língua sem artigo definido	Línguas que possuem artigo definido		
	Forma única	2 formas (masculino/feminino)	3 formas (masc./fem./neutro)
Latim	Inglês: <i>the</i>	Português: <i>o/a</i> Espanhol: <i>el/la</i> Italiano: <i>il/la</i> Francês: <i>le/la</i>	Alemão: <i>der/die/das</i> Grego: <i>ho/hē/tó</i>

Ora, o artigo definido é um termo gramatical, ou seja, que não tem significado lexical (como têm *livro* e *mesa*), mas a função de indicar que aquilo de que se fala é conhecido (o que

se representa como “[+] determinado”)⁶. O fato de o inglês ter apenas uma forma para o artigo definido não interfere nessa função. A única diferença com o português, o espanhol, o italiano, o francês, o alemão e o grego é que, nestas línguas, o artigo concorda em gênero com a palavra que determina.

Agora preste atenção: o latim está bem acompanhado, já que uma grande parte das línguas do mundo não apresenta artigos – e mesmo uma boa parte das línguas indo-europeias, como as eslavas e indo-iranianas. Assim, ‘o livro está sobre a mesa’ se diz, em russo:

КНИГА НА СТОЛЕ.

Veja como encontramos mais um alfabeto, o cirílico, usado pelo russo e por outras línguas eslavas. A frase acima, transliterada para o alfabeto latino, lê-se assim:

KNIGA NA STOLIE
livro sobre mesa
O livro está sobre a mesa.

Para tomarmos mais um exemplo, de uma língua não indo-europeia, vejamos como a mesma frase se diz em turco, da família uralo-altaica (o turco utiliza o alfabeto latino):

KITAP MASA ÜZERİNDE
livro mesa sobre
O livro está sobre a mesa.

De fato, a determinação efetivada pelo artigo definido caracteriza um tipo de línguas bastante restrito e de diferentes famílias, como o basco (língua isolada), o hebraico (língua semítica), o húngaro (língua uralo-altaica) etc. Por outro lado, nem todas as línguas de uma mesma família possuem artigo definido, como no caso do latim.

Agora observe como, no último exemplo, não se usa uma preposição, mas uma posposição – *üzerinde*, ‘sobre, em cima de’. Mas nem sempre preposições ou posposições são necessárias, pois há línguas em que a palavra recebe uma terminação que já expressa diferentes categorias gramaticais, como sujeito, objeto, adjuntos e complementos adverbiais e adnominais – e assim por diante. Isso é o que se chama de sistema de casos. No húngaro, por exemplo, uma palavra pode receber até dezoito terminações para expressar que está na

⁶ Nos exemplos, a presença de artigo determinado, independentemente de sua forma, será indicada pela sigla DET.

função de sujeito (caso nominativo), objeto direto (acusativo), beneficiário da ação (dativo), adjunto ou complemento circunstancial indicando ‘lugar onde’ (inessivo), ‘lugar para onde’ (alativo), ‘lugar de onde’ (ablativo) etc.

Na nossa frase padrão, como se trata de dizer que “o livro está *sobre* a mesa”, a palavra que, em húngaro, significa ‘mesa’, *asztal*, receberá a terminação do caso supressivo (que indica ‘lugar sobre’), assumindo então a forma *asztal-on*, ‘sobre a mesa’:

A KÖNYV AZ ASZTALON VAN
 DET livro DET mesa-SUP⁷ está.

O livro está sobre a mesa.

Como *asztalon* já significa ‘sobre mesa’, estando determinado pelo artigo *az*⁸ – o que faz com que *az asztalon* signifique ‘sobre a mesa’ – não há necessidade de acrescentar uma preposição para indicar ‘sobre’, como nos demais exemplos que vimos até agora.

Agora preste atenção: nas línguas indo-europeias que vimos, as românicas – português, espanhol, francês e italiano – não conhecem flexão de caso e o inglês marca apenas o genitivo, que expressa o possuidor; já o grego, o latim, o alemão e o russo, sim, como se pode constatar abaixo:

Latim

Liber	super	mensa	est.
livro-NOM	sobre	mesa-AC	está.

Russo

Kniga	na	stolie.
livro-NOM	sobre	mesa-PREP

Grego

Tò	biblíon	epì	têi	trapézēi	estí.
DET-NOM	livro-NOM	sobre	DET-DAT	mesa-DAT	está.

⁷ Aos poucos você se acostumará com as siglas que utilizaremos: SUP = supressivo.

⁸ O artigo determinado do húngaro apresenta apenas uma forma (como acontece em inglês): *a*; se, contudo, ele ocorre antes de palavras começadas com vogal, aparece como *az*. Repare no nosso exemplo: *a könyv*, ‘o livro’; *az asztal*, ‘a mesa’.

Alemão					
Das	Buch	ist	auf	dem	Tisch.
DET-NOM	livro	está	sobre	DET-DAT	mesa.

Observe:

- (a) O latim e o russo declinam cada um dos substantivos: 'livro', por ser sujeito das orações, encontra-se no nominativo (que indicamos com a sigla NOM); 'mesa', como complemento circunstancial de lugar, construído com as preposições que significam 'sobre', apresenta-se, em latim, no acusativo (AC), que nesta língua inclui o sentido de locativo ('lugar onde'), e, em russo, no prepositivo (PREP), que tem também a função de expressar o locativo.
- (b) O grego declina tanto os artigos quanto os substantivos, sendo que o locativo se inclui no caso dativo (DAT).
- (c) O alemão declina apenas os artigos, permanecendo os substantivos sem modificação – também nesta língua o dativo (DAT) expressa o locativo.

O mais importante, contudo, é perceber como existe, nas quatro línguas, uma sobrecarga de marcas. Ainda que haja declinação dos nomes e/ou dos artigos, usam-se também preposições.

Uma razão para isso é que, nelas, os casos não são tão específicos quanto no húngaro, em que a terminação *-on* indica apenas 'lugar sobre onde', havendo outros casos para o 'lugar onde', 'lugar para onde' etc. De fato, em grego e alemão, o dativo serve tanto para indicar 'lugar onde', quanto o objeto indireto, dentre outras funções, o mesmo acúmulo de funções variadas acontecendo com o acusativo latino e com o prepositivo do russo. É por isso que as preposições se tornam indispensáveis e são elas que terminam por reger o caso dos nomes com os quais constituem sintagmas.

De um certo modo, essa sobrecarga (preposição + declinação) tende a fazer com que a declinação de artigos e nomes termine por desaparecer em muitas línguas, como ocorreu com as românicas, que procedem do latim, e também com o inglês, que conservou, do germânico, apenas um caso, o genitivo, para indicar o possuidor (*Rose's book*, 'livro da Rose').

* * *

Última observação: você deve ter reparado quantas vezes o imperativo "observe" foi repetido no que você acabou de ler. É que comparar exige isso: treinar a capacidade de observar, para perceber as semelhanças e diferenças. Como nosso tema é a comparação

linguística, então nosso principal objetivo é desenvolver em você essa capacidade de observar o que acontece nas línguas. Uma atitude muito importante para que se torne capaz de observar também o que acontece na nossa própria língua, o português.

Nos textos seguintes, você tomará contato com muitas informações novas e aprenderá umas tantas categorias linguísticas. É evidente que não se espera que você aprenda as tantas línguas a que se fará referência, mas sim – o que é o mais importante – que tome as línguas e a linguística como objeto de conhecimento e de reflexão. Afinal, o homem é um animal que fala, logo, as línguas são um dos traços mais preciosos da condição humana.

Ponto para você que escolheu estudar Letras!

Guia de Leitura Texto 1: O que é Linguística Comparada?

- (1) Por que é possível identificar tantas semelhanças entre as frases escritas nos sete idiomas?
- (2) Identifique os três grupos linguísticos em que o quadro pode ser dividido.
- (3) Qual a origem do termo grego *biblíon*?
- (4) Qual a origem dos termos *book* e *Buch*?
- (5) E da palavra latina *liber*?
- (6) Explique, resumidamente, a série de empréstimos que culminou no termo alemão referente à mesa.
- (7) Além do método histórico-comparativo, qual o outro critério utilizado para classificar as línguas em grupos de semelhança?
- (8) Em quais línguas, elencadas no quadro, o artigo concorda em gênero com a palavra que determina?
- (9) Cite quatro línguas indo-europeias em que se observa a marcação de caso.
- (10) Por que existe uma *sobrecarga de marcas* nestas línguas? Explique.

TEXTO 2

ARQUEOLOGIAS

Jacyntho Lins Brandão

Tanto a diferença quanto a semelhança entre as línguas intrigou desde eras muito antigas a humanidade. É bastante conhecido o episódio da Torre de Babel, através do qual o narrador da Torah⁹, que escreve por volta do século VI a.C., busca dar uma explicação para a diversidade linguística, nestes termos:

Todo mundo se servia de uma mesma língua e das mesmas palavras. Como os homens emigrassem para o oriente, encontraram um vale na terra de Senaar e aí se estabeleceram. Disseram um ao outro: Vamos! Façamos tijolos e cozamo-los ao fogo! O tijolo lhes serviu de pedra e o betume de argamassa. Disseram: Vamos! Construamos uma cidade e uma torre cujo ápice penetre nos céus. Façamo-nos um nome e não sejamos dispersos sobre a terra!

Ora, Iahweh desceu para ver a cidade e a torre que os homens tinham construído. E Iahweh disse: Eis que todos constituem um só povo e falam uma só língua. Isso é o começo de suas iniciativas! Agora, nenhum desígnio será irrealizável para eles. Vamos! Desçamos e confundamos (*nablah*) a sua linguagem para que não mais se entendam uns aos outros. Iahweh os dispersou dali por toda a face da terra, e eles cessaram de construir a cidade. Deu-se-lhe por isso o nome de Babel, pois foi lá que Iahweh confundiu (*balal*) a linguagem de todos os habitantes da terra e foi lá que ele os dispersou sobre toda a face da terra. (*Gênesis*, 11, 1-9. Tradução da Bíblia de Jerusalém, com modificações)

Além da maneira curiosa como a origem da diversidade é apresentada, nada mais que punição pela insolência dos homens, e ainda que a existência de línguas diferentes seja explicada por esse modo, supõe-se que a diversificação aconteceu de chofre, transformando uma situação primitiva quando toda a humanidade falava uma única língua, ou, nas palavras do Rabi Shlomó Yitzkhaki (Rashi, 1040-1105), quando possuía “o bem de ser um só povo com uma só língua”. Não se esclarece, contudo, qual seria essa língua original nem há qualquer traço de que pudesse ser a origem das demais. O que se deseja enfatizar é como a providência tomada por Yahweh, confundindo a linguagem humana, teve o efeito esperado de imediato, ou seja, cessar a construção da torre. Conforme comenta Rashi, na confusão que se instala de

⁹ *Torah* é o nome original que se dá aos cinco primeiros livros da Bíblia judaica, chamados, em grego, Pentateuco. O livro da Torah que, também a partir do grego, conhecemos como *Gênesis*, se chama, em hebraico, *Bereshit*, ou seja, *No princípio*. No Oriente Médio, a partir da prática corrente na Mesopotâmia desde o segundo milênio a.C., era costume que as obras recebessem como título as palavras com que começavam. No caso do *Gênesis*: “No princípio criou Deus o céu e a terra...”.

imediatamente, “um pede um tijolo e o outro lhe traz argila; o primeiro então se enfurece e quebra a cabeça do outro” (YITZJAK, *El Pentateuco*, p. 43-44).

Quase um século após o relato da Torah, encontramos em Heródoto (séc. V a.C.) a descrição da pesquisa levada a cabo por Psamético, faraó do Egito entre 664 e 610 a.C., o qual desejava descobrir que língua e, em consequência, que povo seriam os mais antigos do mundo:

Os egípcios, antes que Psamético os governasse, julgavam que eram anteriores (*prótoi*) a todos os povos. Uma vez que Psamético, quando começou a reinar, quis saber quem seriam os primeiros, disseram-lhe que se pensava que os frígios eram anteriores a eles, egípcios, e eles próprios aos demais povos. Psamético, como não conhecia nenhum meio de descobrir quais seriam os primeiros homens, elaborou este: deu duas crianças recém-nascidas de pessoas de baixa condição a um pastor, para que as alimentasse entre os rebanhos, com o alimento ali usado, ordenando que ninguém, diante delas, emitisse qualquer som (*phonén*); ele devia deixá-las numa cabana solitária e, nos momentos apropriados, levar cabras até elas, dando-lhes leite – e observar o que aconteceria. Psamético fez e levou ao cabo isso por querer ouvir das crianças, quando abandonassem os inarticulados gritos sem significado (*asémon*), qual a primeira palavra (*phonèn próten*) que se poriam a falar. Completados dois anos, ao pastor que cumpria sua tarefa, quando abria a porta e entrava, ambas as crianças, arrastando-se em sua direção, diziam (*ephóneon*) “bekós”, estendendo as mãos. De início, ouvindo isso, ele ficou quieto, mas, como muitas vezes, quando entrava e prestava atenção, essa era a palavra (*épos*), contou-o ao rei. Por ordem deste, conduziu as crianças à sua presença. Tendo-o ouvido o próprio Psamético, informou-se sobre quais dentre homens chamavam algo de “bekós”. Pesquisando, descobriu (*heúriske*) que os frígios assim chamavam o pão. Desse modo, os egípcios aquiesceram, concluindo dessa experiência que os frígios eram mais velhos (*presbytérous*) que eles. (Heródoto, *Histórias* 2, 2. Tradução de Brito Broca, com modificações)

Ressalte-se que esse interesse em saber qual seria a língua primitiva da humanidade não é inocente. Nesse tipo de pensamento, que podemos chamar de *arqueológico*, há três perspectivas culturais entrelaçadas. Num sentido amplo e etimológico, *arqueologia*, palavra composta com os termos gregos *arkhé* e *lógos*, constitui um discurso (*lógos*) sobre o princípio (*arkhé*). Ora, *arkhé* cobre três esferas de significado: (a) a origem no tempo, um começo (como em *arqueolítico*); (b) o ponto de partida de onde outras coisas procedem (como em *arquétipo*); (c) o poder (como em *arconte*, *monarquia*, *oligarquia* etc.). Perguntar, portanto, sobre a origem das línguas envolve os três campos: (a) qual a língua mais antiga? (b) qual a língua donde as demais procedem? (c) qual a língua, que por ser o princípio das demais, exerce sobre elas seu poder e confere poder a quem a conhece? Assim, escolher uma língua qualquer como a original implica atribuir-lhe primazia, em termos de precedência, procedência e poder, supondo-se que aqueles que a falam sejam o povo mais antigo ou descendam diretamente dele, bem como são os detentores da linguagem natural, portanto mais perfeita, de que todas as demais não são mais que devedoras.

Que o assunto manteve seu interesse comprova o fato de que, mais de dois milênios depois, Frederico II, rei do Reino das duas Sicílias e imperador do Sacro Império Romano-Germânico, repetiu, mais de uma vez, a experiência de Psamético, com desfechos fatais:

[Frederico II] quis experimentar qual língua e idioma teriam as crianças, chegando à adolescência, sem terem jamais podido falar com ninguém. E por isso ordenou às amas de leite e às nutrizas que dessem leite aos infantes (...), com a proibição de falar-lhes. Com efeito, queria saber se falariam o hebraico, que foi a primeira língua, ou talvez o grego, ou o latim, ou o árabe; ou se não falariam sempre a língua dos próprios genitores de quem tinham nascido. Mas cansou-se sem resultado, porque as crianças ou infantes morriam todos. (Salimbene da Parma, *Cronaca*, n. 1664, *apud* ECO, 2002, p. 5).

Nesse contexto, há mais um pressuposto importante: o de “língua natural”. As crianças falariam a língua primordial da humanidade (supostamente o hebraico) ou de parcela dela (o grego, o latim ou o árabe, idiomas igualmente antigos) – ou se expressariam na língua materna, ainda que tivessem sido separados das respectivas mães, estando, portanto, impedidos de aprender a falar como todas as crianças? Noutros termos: a língua é inata ou aprendida? Dizendo com mais precisão: é natural ou cultural?

Essa última questão já tinha sido discutida por Platão no *Crátilo* (séc. IV a.C.) e foi enfrentada marginalmente pelo desconhecido autor dos *Discursos duplos* (*Dissoi lógoi*), obra provavelmente escrita no século V a.C. Pela simplicidade como se resolve nesta última obra, mostra-se como é possível encontrar uma resposta sem a necessidade de apelar para experimentos crueis como os de Psamético e Frederico II. O problema que se coloca é se é possível alguém ensinar e aprender. Para solucioná-lo, apela-se para o que se chama de “experiência mental”: dada uma determinada situação, o rigor de análise levará à alternativa correta. Assim, declara o autor:

Se alguém, quando ainda criancinha, fosse mandado para a Pérsia e lá fosse criado, não ouvindo jamais a língua da Grécia, falaria persa; se alguém de lá fosse trazido para cá, falaria grego. (*Dissoi lógoi*, 6, 12)

Portanto, a língua é um dado não da natureza, mas da cultura, e as palavras podem ser ensinadas e aprendidas, uma vez que a criança esteja exposta a determinados contextos, independentemente de sua origem familiar ou étnica.

Observe-se como, nos exemplos citados, há reis dentre aqueles que se preocupam em desvendar qual seria a língua originária da humanidade, o que nos garante a relação entre conhecimento da origem e poder. Não se pense que se trata de uma perspectiva que se perde nas brumas do passado, bastando recordar como o nazismo se apropriou de descobertas no campo da linguística indo-europeia para justificar desmandos e atrocidades, criando o mito da superioridade da raça ariana e de sua pureza (cf. BLIKSTEIN, 1992). Conclusão: trabalhar com a linguagem e as línguas não é algo inócuo ou mera curiosidade, como se poderia pensar.

1 A precedência do hebraico e outras candidaturas

Em geral, a exegese rabínica concordará que aquela “só e mesma língua” utilizada pelos homens no princípio era o hebraico (segundo Rashi, a “língua santa”), ponto de vista adotado também pela quase totalidade dos hermeneutas cristãos antigos e medievais.

Ainda no início da era moderna insistem na mesma tecla, dentre outros, Guillaume Postel (1510-1581) e Claude Duret (1570-1611) – atitude ridicularizada pelo filósofo judeu-holandês Gottfried Leibniz (1646-1716), o qual afirmava que “na suposição de que o hebraico foi a língua original da humanidade há tanta verdade quanto na afirmação do holandês Goropius (...) de que a língua que se falava no Paraíso era justamente o holandês” (NIKOLSKI; JAKOWLEW, 1947, p. 21-22).

A referência de Leibniz é a Goropius Becanus (Jan van Gorp, 1519-1572), modelo de um conjunto mais amplos de autores que defendiam outras candidatas ao posto de língua originária (cada qual puxando a brasa para a própria sardinha): assim, o poeta sueco George Stiernhielm (1598-1672) pretendia que o gótico (ou antigo nórdico) fosse a origem de todas as línguas, assim como os países nórdicos seriam a *vagina gentium*, lugar onde se originou a humanidade; já o médico irlandês James Parsons (1705-1770) opinava que o gaélico era a língua mais próxima da primitiva; e mesmo o filósofo Johann Gottlieb Fichte (1762-1814) defendia que o melhor candidato a língua originária (*Ursprache*) seria o alemão, em vista de sua “pureza”. Outros optaram por soluções mais diversificadas: para um, “Adão falava basco; para outro, ao contrário, Adão e Eva utilizavam o persa, a serpente, que os seduziu, o árabe, e o Arcanjo Gabriel, o turco” (NIKOLSKI; JAKOWLEW, 1947, p. 21-22).

Umberto Eco resume bem os meandros de toda essa pendenga, em que se encontra envolvida a ideia de que a língua original seria também a língua perfeita, o que só comprova como nada se faz por simples curiosidade:

Na sua versão mais antiga, a busca da língua perfeita assume a forma da hipótese monogenética, ou seja, da derivação de todas as línguas de uma única língua-mãe. (...) Os Padres da Igreja, de Orígenes a Agostinho, haviam assumido como um dado incontestável que o hebraico tinha sido, antes da confusão, a língua primordial da humanidade. A exceção mais importante fora a de Gregório de Nissa (*Contra Eunomium*), que sustentara que Deus não falava hebraico e ironizava a imagem de um Deus-professor ensinando o alfabeto a nossos pais. (...) Mas a ideia do hebraico como língua divina sobrevive ao longo de toda Idade Média. Entre os séculos XVI e XVII, não basta mais sustentar que o hebraico era a protolíngua (...): então interessa promover seu estudo e, se possível, sua difusão.

Um lugar particular na história do renascimento do hebraico cabe à figura de erudito utopista que foi Guillaume Postel (1510-1581). (...) No *De originibus seu de Hebraicae linguae et gentis antiquitate* (1538), afirma ele que a língua hebraica provém da descendência de Noé e que dela derivaram o árabe, o caldeu, o índico e, só medianamente, o grego. (...)

Claude Duret, em 1613, publica um monumental *Trésor de l'histoire des langues de cet univers* (...). Já que Duret mantém a ideia de que o hebraico foi a língua universal do gênero humano, é óbvio que o

nome hebraico dos animais contém em si toda sua “história natural”. Assim, “a águia chama-se *neshar*, nome que concorda com *shor* e *isachar*, que significam olhar e estar ereto, porque este pássaro tem, mais que todos, a vista firme e sempre levantada para o sol”. (...)

Mas se Duret fazia etimologia *regressiva*, para mostrar como a língua-mãe estava em harmonia com as coisas, outros farão etimologia *progressiva*, para mostrar como do hebraico derivaram todas as outras línguas. Em 1606, Estienne Guichard escreve *L'harmonie étymologique des langues*, onde demonstra como todas as línguas existentes podem ser reconduzidas a raízes hebraicas. Partindo da afirmação de que o hebraico é a língua mais simples porque nele “todas as palavras são simples e sua substância consiste de apenas três radicais”, elabora um critério que lhe permite jogar com esses radicais por inversões, anagramas, permutações, segundo a melhor tradição cabalística. *Batar* em hebraico significa “dividir”. Como se justifica que de *batar* tenha provindo, em latim, *dividere*? Por inversão se produz *tarab*, de *tarab* se chega ao latim *tribus*, e então a *distribuo* – e a *dividere* (...). *Zaqen* significa “velho”; transpondo-se os radicais, tem-se *zaneq*, donde *senex* em latim; e com uma sucessiva permutação de letras tem-se *cazen*, donde em osco *casnar*, de que derivaria o latino *canus*, que significa justamente “ancião” (...).

O século XVII oferece-nos exemplos saborosos de nacionalismos linguísticos (...). Goropius Becanus (Jan van Gorp), em *Origines Antwerpianae* (1569), sustenta todas as teses correntes sobre a inspiração divina da língua primitiva, sobre a relação entre palavras e coisas, e encontra essa relação exemplarmente presente no holandês, ou melhor, no dialeto de Antuérpia. Os antepassados dos habitantes de Antuérpia, os címbrios, descendem diretamente dos filhos de Jafé, que não se achavam presentes junto da Torre de Babel, escapando assim da *confusio linguarum*. Conservaram, portanto, a língua adâmica, o que se prova através de claras etimologias (...) e pelo fato de que o holandês tem o maior número de palavras monossilábicas, supera todas as outras línguas em riqueza de sons e oferece excepcionais possibilidades para a geração de palavras compostas. (...)

Ao lado da tese holandesa-flamenga não falta a tese “sueca”, com George Stiernhielm (*De linguarum origine praefatio*, 1671). (...)

Quanto ao alemão, várias e repetidas suspeitas sobre seu direito de primogenitura agitam-se no mundo germânico desde o século XIV, em seguida ao pensamento de Lutero (para o qual o alemão é a língua que mais que todas aproxima de Deus), enquanto, em 1533, Konrad Pelicanus (*Commentaria bibliorum*) mostra as evidentes analogias entre alemão e hebraico. (ECO, 2002, p. 83-109)

Enfim, toda essa discussão chegou a tal paroxismo que acabou inteligentemente parodiada pelo filósofo e filólogo sueco Andreas Kempe (1622-1689): em seu panfleto satírico *As línguas do Paraíso* (*Die Sprachen des Paradieses*, de 1688), seu protagonista, Simon Simplex (um Simão simplório qualquer), estabelece que Deus se dirigia a Adão em sueco e este lhe respondia em dinamarquês – enquanto a serpente falava com Eva em francês, já que esta, “a língua tradicional da sedução, ‘mexe com o corpo todo de tal modo que até a pessoa mais sábia pode ser por ela iludida’” (*apud* OSTLER, 2003, p. 1).

2 O que se pode saber sobre a origem da linguagem

Apenas no final do século XVIII e princípios do XIX que o tipo de especulação acima apresentado foi sendo substituído pela ideia de que as línguas do mundo se dividem em diferentes famílias, cujo estabelecimento depende de um paciente trabalho de comparação.

Esse trabalho iniciou-se no âmbito das línguas indo-europeias, num processo paulatino, mas contínuo, marco da fundação da linguística moderna, que abordaremos no capítulo seguinte.

Isso não implica, todavia, que a pergunta sobre a origem – não tanto das línguas, mas da linguagem humana – se tenha tornado improcedente. Com efeito, se toda humanidade tem como característica utilizar-se de línguas como forma de comunicação, isso implica que há, na linguagem humana, um conjunto de categorias universais relacionadas com determinados processos cognitivos, os quais têm recebido cada vez mais a atenção dos linguistas, com enfoques variados.

A diferença com relação às interpretações anteriormente expostas está:

- (a) no estabelecimento de que a língua é um dado de cultura, não da natureza, não havendo, portanto, línguas ou palavras “naturais”;
- (b) no abandono da ideia de que as línguas do mundo possam provir de alguma das línguas conhecidas, uma vez que qualquer língua se encontra em processo de constante mutação;
- (c) na admissão de que é possível, através da comparação, retroceder a estágios anteriores das línguas conhecidas, reconstituindo em parte as protolínguas donde uma determinada família procede;
- (d) finalmente, na constatação de que, a partir da diversidade de línguas e da compreensão de como elas funcionam e se modificam, se podem deduzir certos parâmetros relativos à linguagem humana.

Embora tenha sido abandonada por muito tempo e continue recebendo críticas, a hipótese de que as línguas do mundo tenham uma origem comum voltou a ser considerada seriamente por linguistas como Joseph Greenberg e Merritt Ruhlen, tendo em vista, sobretudo, o avanço do conhecimento relativo às macrofamílias linguísticas, aliado às conquistas da arqueologia, que estuda os dados da cultura material, e, mais recentemente, também da biologia, que vem trabalhando, com bons resultados, no mapeamento do genoma humano. Se o *homo sapiens sapiens* tem uma origem comum – que tudo leva a crer se encontra no continente africano –, é razoável supor que também as várias línguas possam ter uma única origem. Evidentemente, não se poderá jamais saber como seria esse sistema linguístico primeiro, a não ser em termos muito gerais, ou seja, naquilo em que todas as línguas coincidem:

- (a) a arbitrariedade do signo linguístico;
- (b) o uso de categorias linguísticas compatíveis com os processos cognitivos através dos quais o homem apreende o mundo e com ele se relaciona;
- (c) o caráter social da linguagem humana enquanto meio de comunicação;
- (d) o fato de que a língua se encontra em constante processo de variação e mudança.

Considerados esses pressupostos, é legítimo especular sobre a origem da linguagem humana, como faz o linguista alemão Rudi Keller, numa incitante “experiência mental” que, partindo de um dado pressuposto (a evolução da espécie), procura deduzir (sempre hipoteticamente) as várias etapas através das quais, após um gesto inaugural, a linguagem humana se teria separado das formas como os outros animais se comunicam, adquirindo o que tem de específico.

Uma especulação em forma de narrativa (como as míticas), que parece a única possível quando se trata de vislumbrar algo sobre origens que se perdem em tempos imemoriais. É o que você lerá no texto a seguir.

LEITURA COMPLEMENTAR

(KELLER, 1994, p. 19-22)

A história de Chico

Era uma vez um grupo de homens-macaco. Os homens-macaco eram seres que haviam acabado de ultrapassar o estágio de símios, mas não tinham ainda atingido um ponto em que se poderia dizer que eram simplesmente humanos, porque não tinham eles uma linguagem. Todavia, esses homens-macaco tinham a sua disposição, exatamente como seus parentes mais próximos, gorilas e chimpanzés, um rico repertório de expressões sonoras. Os mais coléricos batiam a boca e rosnavam quando estavam irados; os vaidosos batiam no peito e rugiam quando queriam exibir-se. Eles batiam os dentes quando se divertiam, ronronavam quando se sentiam confortáveis e emitiam gritos que rompiam os ouvidos quando ansiosos.

Todas essas manifestações estavam longe de ser signos linguísticos. Não serviam para a comunicação, como hoje a entendemos, mas eram, ao invés disso, a expressão natural de eventos internos: sintomas da vida emocional, comparáveis ao suor, ao frio, ao riso, às lágrimas ou ao rubor. Alguém não comunica suas emoções por meio desses fenômenos, mas, em certas condições, pode revelar algo sobre as mesmas. É que os sintomas podem causar efeitos similares aos dos signos linguísticos.

Um dos integrantes do grupo era um homem-macaco que a natureza pusera em desvantagem: pequeno, mais fraco que os outros e ansioso ao máximo. Podemos chamá-lo de Chico.

Sendo fraco, Chico era muitas vezes forçado, desde a infância, a ser um tanto mais esperto que os outros. Ele tinha de compensar sua falta de força corpórea e seu baixo status social, sob o risco de ficar completamente dominado pelos demais. Em especial, os membros mais fortes do grupo afastavam-no regularmente da comida, não deixando que ficasse perto dos bocados mais suculentos. Mas, sendo ágil e esperto, Chico conseguia ultrapassar alguns desses obstáculos.

Um dia aconteceu algo que teria uma imensa importância para o futuro de toda raça dos homens-macaco. O grupo estava pacificamente amontoado em volta da comida, consumindo a presa capturada naquele dia. Como sempre, havia algumas pequenas brigas e empurrões ocasionais. Chico foi de novo empurrado para a borda exterior, onde descobriu um par de olhos no meio da vegetação rasteira – os olhos de um tigre! Seus olhos encontraram-se com os do animal... Morrendo de medo, ele grita aterrorizado. O grupo dispersa instantaneamente. Cada qual trata de encontrar abrigo na árvore mais próxima, porque tal grito era sinal de enorme perigo. Estavam todos condicionados, desde a infância, a reagir assim.

Chico ficou parado lá, como se congelado. Estar tão perto de morrer o havia tornado incapaz de fugir. Todavia, para seu grande espanto, os olhos piscavam para ele, de um modo nada parecido com o que faz um tigre, e seu proprietário foi-se embora irritado. O que ele havia visto como olhos de tigre pertencia a nada mais que um pacato porco. Chico tinha sido vítima de sua vívida imaginação, alimentada por sua natural ansiedade.

Mas “vítima” é a palavra correta neste caso?

Quando Chico olhou em volta, desconcertado, desamparado e um pouco envergonhado, viu que estava completamente sozinho, junto com a comida deixada para trás pelos outros. A expressão de medo em seu rosto deu lugar a um firme e travesso sorriso. Ele quase não podia acreditar.

Na medida em que passavam os dias e as semanas – e que, a cada vez, a disputa pelas melhores partes de alimento tinha lugar – ele era tentado a fazer intencionalmente o que lhe havia acontecido por acidente. O que Chico não podia imaginar é que essa tentação marcava o fim do paraíso da comunicação natural.

O que tinha de acontecer finalmente aconteceu. Como sempre, ele tinha de ficar observando como aqueles grandalhões cabeludos repartiam as melhores partes entre si, enquanto ele, faminto, se sentava perto, tomado por uma raiva impotente. Foi então que sucumbiu à tentação. Deu o grito de angústia e, de novo, o grupo dispersou-se em matéria de segundos, incluindo os repugnantes grandalhões.

A melhor parte da comida ficara ali, montes de comida. Na sua agitação, Chico, na verdade, nem pôde saboreá-la (talvez sua má consciência o impedisse). Mas o primeiro degrau tinha sido galgado e Chico achou muito mais fácil da próxima vez. Com o tempo, tornou-se quase impiedoso. Achava prazer em executar seu truque e começou mesmo a abusar.

Como era inevitável, logo alguém suspeitou dele. Quando Chico foi bobo o suficiente para gritar pela segunda vez durante uma mesma tarde, um outro macaco parou, depois de poucos saltos, olhou para trás e começou a devorar a comida. Chico ficou um pouco irritado, mas não se incomodou, pois havia comida suficiente para ambos. Mas logo o cúmplice começou também a usar do expediente que aprendera e, como Chico, a exagerar.

O número daqueles que não se deixavam enganar pela mentira – e, finalmente, o número de imitadores – tomou dimensões inflacionárias. A comunidade entrou num período extremamente crítico. Cada qual suspeitava dos demais. Os grandalhões tentaram restaurar a

antiga ordem, penalizando todo abuso do grito de prevenção. Mas um conhecimento, uma vez adquirido, não pode ser jamais erradicado. Pelo contrário, era reforçado por todo novo abuso e toda tentativa de penalizar quem dele utilizava.

O abuso permanente do grito de prevenção representava um perigo para a existência física de todo o grupo, uma vez que a crença cega nele era necessária para a sobrevivência. Mas essa época havia definitivamente acabado. Os que quisessem sobreviver nesses tempos de corrupção tinham de ter bons ouvidos. Tinha de aprender a diferenciar o grito genuíno do fingido, algo que não se mostrou difícil para muitos deles. (...)

A história de Chico não pretende ser realista, mas diz algo sobre a realidade. Ela mostra como a transição da comunicação natural para a humana *poderia* ter acontecido. Não se trata de uma reconstrução histórica, mas filosófica. Não são os fatos, mas apenas os dados lógicos da história que devem estar corretos, a saber:

1. As etapas que conduzem do grito natural de angústia ao ato intencional parecem plausíveis. A passagem de um ao outro não deve ter apresentado nem furos, nem saltos.
2. As pressuposições relativas às habilidades dos homens-macaco parecem ser realistas. A história seria sem valor caso se atribuísse a Chico uma alta (e irrealista) capacidade intelectual.

EXERCÍCIO

Tomando como base a história de Chico, discuta os seguintes aspectos:

- a) Qual a diferença entre sintoma e linguagem?
- b) Quando se pode dizer que a arbitrariedade do signo linguístico se manifesta, criando a linguagem humana?
- c) Que papel tem nisso a “mentira” – ou, se quisermos, a capacidade de “fingir”, isto é, a “ficção” como uma função básica da linguagem?
- d) Através de quais processos se manifesta, na história de Chico, o caráter social da linguagem?
- e) A história de Chico dá a entender que a linguagem humana está na ordem da natureza ou da cultura?
- f) Você concorda que a linguagem humana deva ser considerada, como outros, “um conhecimento” que, “uma vez adquirido, não pode ser jamais erradicado”?

Guia de Leitura Texto 2: Arqueologias

- (1) Quais as três esferas de significado cobertas pelo termo *arkhé*?
- (2) Qual o resultado da experiência de Frederico da Prússia, inspirada no precedente estabelecido na Antiguidade por Psamético?
- (3) Explique porque a língua não é um dado da natureza.
- (4) Além do hebraico, quais línguas pleitearam a posição de língua original?
- (5) Explique, resumidamente, qual a tese sustentada por Goropius Becanus com relação à língua adâmica.
- (6) Como o filósofo e filólogo Andreas Kempe parodiou essa tese?
- (7) Com relação à noção de uma suposta língua originária, quais as quatro principais diferenças de abordagem entre a lingüística moderna e os filólogos e estudiosos anteriores a ela?
- (8) A possibilidade de uma língua original tem sido defendida por linguistas como Joseph Greenberg e Merrit Ruhlen. Quais evidências e indícios parecem corroborar a tese destes linguistas?
- (9) Cite quatro características comuns a todas as línguas.

TEXTO 3

O MÉTODO HISTÓRICO-COMPARATIVO

Jacyntho Lins Brandão e Júlio César Vitorino

O método comparativo aplicado ao estudo das línguas constitui uma área de conhecimento que conta mais de duzentos anos. Com efeito, o comparatismo situa-se na base da formação da linguística moderna, a qual tem início, no século XIX, com o estudo das línguas indo-europeias, a então chamada “gramática comparada”, que visava à reconstituição da “língua-mãe”. Nesse domínio, em que se lidava com um conjunto de fatos relativamente limitado, envolvendo, de início, a comparação entre o grego, o latim e o sânscrito, a que logo se juntam o gótico e o celta, teorias e métodos puderam ser testados, na esfera da fonética, da fonologia, da morfologia e, em escala menor, também da sintaxe.

Apenas para citar um exemplo, Ferdinand de Saussure (1857-1913) – cujo *Curso de Linguística geral*, publicado postumamente, em 1916, a partir das anotações de aula feitas por seus alunos, é considerado a primeira obra da linguística moderna – formou-se no campo da linguística comparada, tendo-se tornado conhecido pela *Memória sobre o sistema primitivo das vogais nas línguas indo-europeias*, livro aparecido em 1879.

1 A descoberta do indo-europeu

Pode-se afirmar que o fato mais determinante para a fundação da linguística comparada foi a “descoberta” do sânscrito pelos estudiosos europeus, o que se dá a partir do século XVI, num processo que se desdobra lentamente. De fato, apenas no século XVIII surgiram a primeira gramática sânscrita escrita por um europeu e o primeiro dicionário malaio-sânscrito-português (o português era, então, a língua europeia predominante na Índia), ambos da autoria do jesuíta alemão Johann Ernst Hanxleden (1681-1732). Contudo, a obra que foi fundamental para difundir o sânscrito na Europa só aparece em 1808, da autoria de Friedrich Schlegel (1772-1829), *Über die Sprache und die Weisheit der Inder (Sobre a língua e a sabedoria dos hindus)*, a partir da qual ganhou força a tese da existência de um grupo linguístico indo-europeu.

O conhecimento do sânscrito foi importante para os estudos comparativos tendo em vista, principalmente, três fatores: (a) não se tratava mais de descobrir qual teria sido a língua primitiva da humanidade; (b) os claros paralelismos entre, de um lado, o sânscrito e, de outro, o grego e o latim não poderiam ser explicados pelo contato entre esses povos em qualquer período conhecido da história; (c) pouco a pouco, a constatação das semelhanças sobretudo entre essas três línguas conduziu à formulação da hipótese de que deveriam elas pertencer a uma mesma “família”. Comparação e “parentesco” linguístico passam a ser, portanto, as palavras de ordem.

É bem verdade que, ainda na Antiguidade, a semelhança entre o grego e o latim intrigara os romanos, o que não seria para menos, bastando observar algumas das palavras para cuja correspondência chamou atenção o gramático Varrão (116 a.C.-27 d.C.) em seus estudos de “etimologia” (cf. *Da língua latina*, fr. 5, 34 e 96):

QUADRO 1
Comparação de termos gregos e latinos

Significado	Grego	Latim
campo	<i>agrós</i>	<i>ager</i>
porco	<i>hús</i>	<i>sus</i>
boi	<i>boús</i>	<i>bos</i>
touro	<i>taûros</i>	<i>taurus</i>
ovelha	<i>óis</i>	<i>ouis</i>

Como gregos e romanos eram povos geograficamente vizinhos, a explicação mais corrente para tais correspondências foi considerar que o latim derivava do grego, mais especificamente de um de seus dialetos, o eólico. Essa opinião, que tinha antecedentes em autores romanos como Catão o Velho (234-149 a.C.), foi assim expressa pelo historiador grego Dionísio de Halicarnasso (60-7 a.C.): “Os romanos falam uma língua que não é nem totalmente bárbara, nem completamente a grega, mas uma certa mistura de ambas, de que a maior parte é eólica” (*Antiguidades romanas* 1, 90, 1). De novo é preciso lembrar que nada se diz sem intenção: o que se pretendia, neste caso, era não tanto tratar das línguas, mas enobrecer o povo romano, atribuindo-lhe ancestrais gregos (cf. GABBA, 2000).

Gregos e latinos, nem na Antiguidade, nem na Idade Média, avançaram além desse nível superficial de comparação. Somente no século XVI, com as grandes navegações, quando os primeiros europeus entraram em contato com as línguas de outros continentes, a história começaria a tomar outros rumos. Também neste caso, é preciso assinalar, o interesse da catequização e da exploração das colônias fundadas na África, na América e no extremo Oriente foram os propulsores do interesse pelas línguas. Recorde-se que é dessa época, por exemplo, a *Gramática da língua geral da costa do Brasil*, do padre José de Anchieta. Neste caso e

em tantos outros tratava-se de produzir obras com finalidades bastante pragmáticas. Mas foi a uma parte das línguas da Índia, aquelas pertencentes ao grupo indo-europeu, em especial ao sânscrito, que caberia um papel decisivo, por permitir comparações cada vez mais detalhadas com uma parte das línguas da Europa.

Já em 1583, numa carta que permaneceu inédita até o século XX, o jesuíta inglês Thomas Stephens (1549-1619) dava notícia da existência de paralelos entre o concâni, o grego e o latim. Dois anos mais tarde, em 1585, o mercador florentino Filippo Sassetti (1540-1588), que viveu algum tempo em Goa e estudou sânscrito, chamou a atenção para a correspondência entre esta língua e o italiano, como, por exemplo, nas palavras abaixo:

QUADRO 2
Comparação de termos italianos e sânscritos

Significado	Italiano	Sânscrito
sete	<i>sette</i>	<i>sapta</i>
nove	<i>nove</i>	<i>nava</i>
deus	<i>dio</i>	<i>deva^h</i>
serpente	<i>serpe</i>	<i>sarpa^h</i>

Em 1768, o jesuíta francês Gaston Coeurdoux (1691-1779) apresenta à *Académie des Inscriptions et les Belles Lettres*, em Paris, estudos comparativos sobre o sânscrito, o grego e o latim, que incluíam minucioso confronto das conjugações verbais nas três línguas, propondo que as semelhanças só se podiam dever a uma origem comum. Entretanto, como seus trabalhos só foram publicados em 1808, sua tese não encontrou imediata ressonância.

A “descoberta” oficial e o início dos estudos de linguística indo-europeia têm sua data emblemática: a comunicação do diplomata inglês William Jones (1746-1794) à Real Sociedade Asiática de Calcutá, em 1786, quando ele avança a hipótese de que sânscrito, grego, latim e – ele apenas supõe então – também o gótico, o celta e o persa provêm de uma mesma origem. Seu objeto não é abordar a questão da proximidade entre essas línguas, mas tão somente tratar da antiguidade dos povos indianos, examinando, “em primeiro lugar, as suas línguas e escritas; em segundo lugar, a sua filosofia e religião; em terceiro lugar, os restos atuais de sua antiga escultura e arquitetura, os memoriais escritos de suas ciências e artes”, estendendo a comparação a todos esses domínios. Assim, por exemplo, ele aproxima Apolo de Kṛshna, afirmando ainda que “não é possível ler o Védânta ou as várias refinadas composições que o ilustram, sem crer que Pitágoras e Platão derivaram suas sublimes teorias da mesma fonte que os sábios da Índia” (JONES, 2009, p. 19-34).

O trecho que se tornou famoso para os estudos linguísticos, portanto, não constitui mais que uma observação de passagem, em que o objetivo principal é ressaltar as características excepcionais que ele, Jones, percebia no sânscrito:

A língua sânscrita, seja qual for a sua antiguidade, tem uma estrutura admirável; mais perfeita que o grego, mais copiosa que o latim e mais elegantemente refinada que cada um deles, ainda que mantenha com ambos, seja nas raízes dos verbos, seja nas formas gramaticais, uma afinidade mais forte do que seria possível produzir-se por acidente – deveras tão forte que nenhum filólogo poderia examinar todas as três sem crer que tenham advindo de alguma fonte comum, a qual talvez não exista mais há muito tempo; há uma razão similar, embora não absolutamente tão impositiva, para supor que também o gótico e o celta, embora mesclados com um idioma bastante diferente, tenham a mesma origem que o sânscrito; e o antigo persa deveria ser adicionado à mesma família, se houvesse ocasião para discutir alguma questão relativa às antiguidades da Pérsia. (JONES, 2009, p. 19-34)

Ainda que com toda essa concisão, na verdade a declaração de Jones teve como mérito:

(a) ressaltar a proximidade entre sânscrito, grego, latim, germânico, celta e persa (ou seja, representantes de cinco dos dez grupos de línguas indo-europeias hoje admitidos); (b) não imaginar que uma das línguas conhecidas fosse a origem das demais, mas postular que deveriam elas provir de uma fonte comum (*some common source*) talvez não mais existente (*which, perhaps, no longer exists*); (c) isso posto, atribuir as afinidades ou parentesco (*affinity*) ao fato de que todas essas línguas deveriam pertencer a uma mesma família (*the same family*). Mais que tudo, observe-se, no uso dos condicionais e dos advérbios, que não se trata de formular postulados, mas de levantar hipóteses cuja comprovação dependeria de outros estudos. Como observa Blikstein, tendo em vista os antecedentes acima expostos,

na verdade, as semelhanças entre o sânscrito e as línguas europeias já tinham sido percebidas bem antes do séc. XIX. Ocorre, no entanto, que a história das ideias e do pensamento não é linear; ao contrário, ela é descontínua e, no dizer do eminente linguista romeno Eugenio Coseriu (...), a história da ciência linguística é “cheia de ocos, a tal ponto que, reiteradamente, as mesmas coisas voltam a ser ‘redescobertas’”. (BLIKSTEIN, 1992, p. 105)

O passo seguinte da “redescoberta” será dado pelo livro de Schlegel já referido, mas o mais importante foi a publicação, em 1816, do estudo do alemão Franz Bopp (1791-1867) *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jener der griechischen, lateinischen, persischen, und germanischen Sprache* (Sobre o sistema de conjugação da língua sânscrita em comparação com o das línguas grega, latina, persa e germânica). O estudo de um sistema, neste caso o verbal, revelou-se um argumento mais poderoso para fundamentar a hipótese de uma origem comum que a simples comparação lexical. Estava, portanto, criado o método comparativo, cujo pressuposto de partida é que

entre elementos de línguas aparentadas existem correspondências sistemáticas (e não apenas aleatórias) em termos de estrutura gramatical, correspondências estas passíveis de serem estabelecidas por meio duma cuidadosa comparação. Com isso, podemos não só explicitar o parentesco entre línguas (isto é, dizer se uma língua pertence ou não a uma determinada família), como também determinar, por inferência, características da língua ascendente comum de um certo conjunto de línguas. (FARACO, 2005, p. 134).

Na sequência, entre 1833 e 1852, o próprio Bopp estendeu a comparação ao lituano, eslavo, armênio, celta e albanês, abrangendo, assim, todos os grupos indo-europeus então conhecidos (ainda não tinham sido decifrados nem o hitita nem o tocário). Paralelamente, já

em 1818 havia aparecido o trabalho do dinamarquês Ramus Rask (1787-1832), em que eram comparadas as línguas nórdicas com as demais línguas germânicas, o grego, o latim, o lituano, o eslavo e o armênio, o que só confirmava a validade de se considerar que todos eles integravam uma mesma família.

Você mesmo poderá observar, nos exemplos abaixo, o quanto essa hipótese é plausível. Vamos examinar os nomes de parentesco, um grupo de palavras que tende a ser bastante conservador. O quadro 3 apresenta quatro línguas antigas e não mais faladas (avéstico, sânscrito, grego clássico e latim) e quatro línguas modernas e faladas hoje em dia (irlandês, alemão, inglês e russo).

QUADRO 3

Nomes de parentesco nas línguas indo-europeias

Sentido	Avéstico	Sânscrito	Grego	Latim	Irlandês	Alemão	Inglês	Russo
mãe	<i>matár-</i>	<i>matar</i>	<i>méter</i>	<i>mater</i>	<i>máthair</i>	<i>Mutter</i>	<i>mother</i>	<i>mati</i>
pai	<i>pitár-</i>	<i>pitar</i>	<i>patér</i>	<i>pater</i>	<i>athair</i>	<i>Vater</i>	<i>father</i>	<i>otiets</i>
irmão	<i>bhrátar-</i>	<i>bratar-</i>	<i>adelphós</i>	<i>frater</i>	<i>bráthair</i>	<i>Bruder</i>	<i>brother</i>	<i>brat</i>
filha	<i>duhitár-</i>	<i>dugadar-</i>	<i>thugáter</i>	<i>filia</i>	<i>iníon</i>	<i>Tochter</i>	<i>daughter</i>	<i>dotsiernii</i>
irmã	<i>svásar-</i>	<i>x^var̥har-</i>	<i>adelphé</i>	<i>soror</i>	<i>siúr</i>	<i>Schwester</i>	<i>sister</i>	<i>siestra</i>
filho	<i>sunuh</i>	<i>hunuš</i>	<i>huiós</i>	<i>filius</i>	<i>mac</i>	<i>Sohn</i>	<i>son</i>	<i>syn</i>

A primeira constatação é que nem todas as palavras no quadro têm a mesma origem, algumas línguas tendo procedido a substituição do termo que seria comum. Contudo, a grande maioria conserva denominações muito próximas. Para descobrir quais vêm de um étimo comum, faça assim:

- Antes de tudo, tenha em vista os fonemas que têm alguma proximidade e procure observar se há alguma regularidade nas correspondências entre eles: as oclusivas dentais *t/th/d/dh*; as oclusivas velares *k/kh/g/gh* e a fricativa *h*, que pode derivar das suas formas aspiradas; as oclusivas labiais ‘*p/ph/b/bh*’ e as fricativas que podem derivar de suas formas aspiradas, ou seja, *f/v*; finalmente, as fricativas *s/x* e a aspirada que pode derivar da primeira, *h*.
- Agora comece por ‘mãe’, que apresenta semelhanças bastante regulares em todas as línguas: você constatará que todos os termos são compostos de duas sílabas (a primeira *-ma/me/mo/mu-*, seguida de *-tar/thair/ter/ther/ti*), o que sugere que todas devem proceder do mesmo étimo, reconstituído inicialmente como **māter*;
- Passe em seguida para ‘pai’, em que a mesma terminação se repete, ocorrendo na primeira sílaba as variações *pa/pi/fa/va-* (que podem ser explicadas considerando o referido no item ‘a’ acima), a forma do irlandês apresentando a seguinte evolução: **pa- > pha- > ha- > a-*; em conclusão, a fonte de todos os termos (com exceção da

palavra russa, que tem origem diferente) foi reconstituída tradicionalmente como **pater*.

- d) Em 'irmão' você perceberá que o grego apresenta uma palavra de origem diferente, podendo todas as demais ser explicadas a partir do que você já conhece, o étimo tendo sido primeiramente reconstituído como **bhrater*.
- e) Os termos correspondentes a 'filha' apresentam a mesma formação que os demais (com o sufixo **-ter*), podendo as diferenças ser explicadas a partir do que você já conhece; o étimo comum foi em princípio reconstituído como **dhughter*, as exceções ficando por conta do latim, do irlandês e, parcialmente, do russo (o antigo eslavo eclesiástico apresenta a forma *dušti*).
- f) Com relação a 'irmã', a exceção se deve de novo ao grego, as demais palavras podendo ser aproximadas e sua origem tendo sido reconstituída de início como **swesor* (no latim, a mudança *s > r* é normal em certos contextos).
- g) Finalmente, para 'filho' as exceções se constatarem no latim e no irlandês, para os demais termos tendo sido proposto inicialmente o étimo **sunu-*.

Todas essas reconstituições representam tentativas iniciais que mais tarde conhecerão outras propostas de restabelecimento, na medida em que se diferenciam as formas como se desenha o sistema fonológico do indo-europeu. Em especial, o vocalismo apresentava problemas que só aos poucos foram esclarecidos. Mas alguns elementos se impõem de imediato, como o uso de **-ter* para marcar uma parte desses nomes de parentesco, o que leva a supor que se trate de um sufixo (cuja produtividade parece ter sido estendida, no germânico e no russo, além dos quatro primeiros nomes do quadro, abrangendo também o relativo a 'filha').

De qualquer modo, é evidente que a semelhança entre tantos termos com os mesmos significados não se poderia dever a mero acaso – ou seja, o único modo de buscar uma explicação razoável é apelar para a hipótese de uma origem comum de todas essas línguas.

2 A constituição do método histórico-comparativo

Uma etapa importante para a constituição do arcabouço da chamada *gramática* ou *linguística histórico-comparativa* foi a publicação, em 1819, da *Deutsche Grammatik (Gramática alemã)*, por Jacob Grimm (1785-1863) – um dos dois irmãos que ficaram famosos por terem recolhido e publicado os contos de fadas (*Märchen*) da tradição alemã. A diferença entre o trabalho de Grimm e o de seus antecessores, Bopp e Rask, está no fato de que ele não estava interessado apenas em comparação visando à demonstração do parentesco entre línguas, mas desenvolveu um estudo propriamente histórico, abordando o desenvolvimento do grupo

linguístico germânico durante quatorze séculos e atento a mudanças cronologicamente datáveis.

Uma das “descobertas” mais importantes de Grimm foi a lei que passou a ser chamada por seu nome (“lei de Grimm”), segundo a qual se mostrava, com relação às consoantes oclusivas, o caráter regular de mudanças que separavam o ramo germânico dos demais grupos indo-europeus (fenômeno também chamado de “primeira rotação consonantal germânica”): (a) onde, no indo-europeu, se encontrava uma oclusiva surda ou desvozeada (*p, t, k*), no germânico se tinha uma aspirada ou uma fricativa dela derivada (*f, th, h*); (b) onde, no indo-europeu, havia uma aspirada (*ph, th, h*), o resultado era uma sonora ou vozeada (*b, d, g*); (c) finalmente, onde, no indo-europeu, existia uma vozeada (*b, d, g*), em germânico se encontraria uma desvozeada (*p, t, k*). Assim, por exemplo:

QUADRO 4
Lei de Grimm

desvozeada > aspirada i.e. * <i>trei-</i>		aspirada > vozeada i.e. * <i>bhrater</i>		vozeada > desvozeada i.e. * <i>dwo-</i>	
latim: <i>tres</i>	gótico: <i>threis</i>	sâncsc.: <i>bhratar-</i>	gótico: <i>brothar</i>	latim: <i>duo</i>	gótico: <i>twa</i>
grego: <i>treis</i>	inglês: <i>three</i>	latim: <i>frater</i>	inglês: <i>brother</i>	grego: <i>duo</i>	inglês: <i>two</i>

Outra contribuição importante para os estudos histórico-comparativos deu-se na esfera de um grupo bastante conhecido das línguas indo-europeias: o românico. Entre 1836 e 1844, Friedrich Diez (1794-1876) publicou trabalhos histórico-comparativos das línguas procedentes do latim, incluindo um dicionário etimológico das mesmas (1854), o que marca a criação da Filologia Românica, uma disciplina que conheceu, a partir de então, um desenvolvimento extenso e notável. Neste caso, havia uma grande vantagem: podia-se contar com registros escritos da língua de origem – o latim – e das dela derivadas: português, galego, espanhol, catalão, provençal, francês, retorromano, italiano, sardo e romeno. Se, com relação a outros grupos, a reconstituição da protolíngua ficava no campo das hipóteses, com a Filologia Românica a validade dos métodos de comparação pôde se testada.

Vale lembrar um dos casos. Pelo método comparativo, aos poucos foram estabelecidos os processos de mudança fonética do latim para as diferentes línguas românicas. Um dos exemplos seria o grupo inicial latino *pl-*, que, em certas condições, gera os seguintes resultados: *ch-* em português; *ll-* em espanhol; *pl-* em francês; *pi-* em italiano. Uma série de termos confirma essa regra:

QUADRO 5
O grupo inicial *pl-* em latim e em línguas românicas

português	espanhol	francês	italiano	latim
<i>chão</i>	<i>llano</i>	<i>plain</i>	<i>piano</i>	<i>planu-</i>
<i>cheio</i>	<i>lleno</i>	<i>plein</i>	<i>pieno</i>	<i>plēnu-</i>

Acontece que, com relação a um dos casos, o do verbo ‘chover’, a previsão se aplicava perfeitamente com relação ao grupo consonantal inicial, mas observava-se um problema no final da palavra:

QUADRO 6

Chover

português	espanhol	francês	italiano	latim
<i>chover</i>	<i>llover</i>	<i>pleuvoir</i>	<i>piovere</i>	<i>*plouëre</i>

Ora, se a comparação, considerados os processos de mudança fonética, levava a *plouëre*, a documentação garantia que, no latim clássico, o verbo ‘chover’ era *pluëre*, motivo por que a palavra foi marcada com um asterisco, para indicar que se tratava apenas de um termo reconstituído, mas não atestado. Com efeito, a semivogal latina *-u-* deveria gerar *-v-* nas línguas relacionadas, bem como os verbos latinos da terceira conjugação (terminados em *-ēre*) resultam regularmente nas terminações *-er/-oir/-ere* nessas mesmas línguas (cf. lat. *mouēre* > port. e esp. *mover*, fr. *mouvoir*, it. *muovere*).

Todavia, mais tarde se “descobriu”, num episódio famoso de um romance escrito no séc. I ou II d.C., o *Satiricon* de Petronio, o termo então apenas reconstituído: durante a ceia de Trimalquião, um novo rico, várias personagens contam histórias e o autor, por tratar-se de um texto satírico, reproduz muitos elementos da linguagem comum; numa dessas histórias, devida ao personagem Ganimedes, num determinado momento se declara que *urceatim plouebat* (chovia a cântaros, *Satiricon* 44). Ora, o imperfeito de *pluëre* é *pluebat*, sendo a forma *plouebat* própria do que se passou a chamar de latim vulgar, ou seja, o latim falado, de onde procedem efetivamente as línguas românicas. O curioso é que, embora *plouebat* fosse atestado nos manuscritos do *Satiricon*, os editores corrigiam o termo, julgando que se devesse a confusão de algum copista medieval – e, com efeito, anteriormente à reconstituição da passagem pelos romanistas do século XIX, o que se conhecia era apenas a forma corrigida, *urceatim pluebat*, como, por exemplo, ela é citada por Voltaire no seu *Dictionnaire philosophique* (*Dicionário filosófico*) de 1764 (verbete *Idole, idolâtre, idolâtrie*, seção I). Portanto, observe como, a par do fato de que esse episódio confirma a validade do método comparativo, mostra também como os estudos histórico-comparativos induziram a um cuidado maior na busca e leitura de fontes capazes de fornecer informações sobre as mudanças linguísticas.

Com o correr dos anos, outros grupos de línguas indo-europeias foram também objeto de estudos específicos – como as eslavas, as celtas, as indo-iranianas etc –, o que terminou por fornecer aos estudos histórico-comparativos uma boa base de dados empíricos, indispensável para testar sua metodologia.

3 Teorizações sobre a linguagem

Uma das consequências mais importantes dos estudos comparados das línguas indo-europeias foi ter induzido a teorizações sobre a língua e a linguagem, fundando a linguística moderna.

É dos passos desse percurso que você encontrará um apanhado no texto complementar que você lerá a seguir.

LEITURA COMPLEMENTAR

Rita D'Avino

Introduzione a un corso di storia comparata delle lingue classiche, 1997, p. 3-17

História da linguística histórico-comparativa

A história da língua como objeto de pesquisa científica é uma conquista relativamente recente, tornada possível pela profunda transformação da cultura denominada "Romantismo". Nela, a exaltação da fisionomia particular dos povos se traduz na busca dos traços que tais fisionomias individualizam: a língua, os costumes, o direito. A noção de história linguística permanece até então improdutiva no plano concreto da pesquisa, ainda que intuída em várias retomadas e em medida diversa desde a Idade Média, e mesmo se teorizada no quadro de temáticas filosóficas mais amplas.

Na Antiguidade, gregos e latinos observaram a língua com interesses filosóficos, lógicos, literários, normativos, enquanto ignoraram o problema histórico. Assim, aos antigos indianos faltou completamente a noção do desenvolvimento histórico, mesmo tendo feito de *Vac* uma divindade, reconhecendo nela a nominabilidade das coisas inerente às próprias coisas e mesmo tendo-se dedicado com profunda atenção à observação da sua língua a todos os níveis (fonetismo, morfologia, sintaxe, estilo). E eles não tiraram partido, para uma reflexão comparativa, do antigo iraniano, nem do grego, quando as vicissitudes históricas os colocaram em contato com esses povos. Analogamente, os gregos, mesmo que sejam considerados não injustamente os fundadores da gramática¹⁰, não trouxeram interesse histórico aos fatos linguísticos e consideraram *bárbaroi* as linguagens de tantos povos (persas, frígios, armênios, trácios, ilíricos) com os quais tiveram contato. O fato que, por exemplo, Heródoto ofereça anotações agudas sobre características do iônico o do dórico de Siracusa e de Cirene, sobre certas parentelas linguísticas ou sobre relações de vocábulos gregos com línguas bárbaras apenas confirma que observações desse tipo, se não faltaram de todo, não encontraram, porém, o terreno apto a transformar a descrição de fatos esporádicos em uma análise metódica do seu desenvolvimento.

Assim, por toda a Idade Média, o Renascimento, até o início do século dezenove, não obstante intuições penetrantes e, como foi dito, precisas aquisições teóricas: de Dante, que vê no *ydiuma trifarium* a comunhão linguística do *sì*, *oc* e *oil*, aos gramáticos italianos do século XVI (Cláudio Tolomei, Celso Cittadini, que começam a observar de perto as relações entre as línguas românicas vulgares e o latim; aos exegetas bíblicos que, comparando as três redações

¹⁰ É sabido que a Aristóteles remonta a individuação das categorias gramaticais, a divisão tradicional das partes do discurso, a terminologia; que estoicos e peripatéticos constituíram aquele complexo de doutrinas gramaticais que, assimiladas e elaboradas pelos latinos formaram a *ars grammatica* antiga e medieval, da qual depende a reflexão gramatical até o surgimento da linguística moderna; que os alexandrinos estudaram e descreveram nas suas específicas qualidades dialetais os textos literários)

(hebraica, grega e latina) da Bíblia (ainda que com intentos teológicos) chegam à noção de “gênio da língua”; até Bacon, Locke, seguido pelos outros empiristas ingleses, a Leibniz e, independentemente, a Giambattista Vico, cuja voz isolada se opõe ao racionalismo, que imperava na cultura do seu tempo, com o *De causis linguae latinae* de Giulio Cesare Scaligero e a *Minerva* de Francesco Sanzio: a estes ele repreendeu por procederem “como se os povos que encontraram as línguas tivessem antes de ter andado a escola de Aristóteles, com cujos princípios ambos raciocinavam”. Em particular, a novidade revolucionária de Vico – com o reconhecimento que o signo linguístico nasce como fato intuitivo e fantástico, e não como qualificação racional do real – foi uma afirmação explícita e consciente da historicidade do fato linguístico (as línguas mudam por intervenção de momentos poéticos no tecido dos “falares convencionais”). Todavia, quase dois séculos após esta concepção pode se afirmar sobre o plano dos estudos linguísticos, em que múltiplas observações e intuições, às vezes felizes, desde a Idade Média, mas especialmente nos séculos XVI, XVII e XVIII, com sensível progresso nesse último, surgiram do confronto e da tentativa de agrupamento dos materiais linguísticos, afirmando-se, exatamente, quando os tempos foram culturalmente maduros para exprimi-los e torna-los produtivos.

Concorreu para isso, sem dúvida, a impositação cultural mencionada, na qual, com novo interesse se observou a história de cada um dos povos, o gosto tipicamente romântico. É pelo gosto por tudo o que é exótico ou primitivo, ou de qualquer modo distante no tempo e no espaço que se deve a ressonância de um livro como o de Friedrich Schlegel, *Über die Weisheit und Sprache der Indiers* (1808). A comparação com o antigo indiano é o momento iniciador da linguística como ciência. Schlegel nota a afinidade do sânscrito com o latim, o grego, o germânico, o persa nos seguintes termos: “a afinidade reside não só no grande número de raízes que ele (o sânscrito) tem em comum com essas línguas, mas se estende também à estrutura mais interna e à gramática. A concordância, portanto, não é casual e tal que se possa explicar por misturas, é concordância substancial, que evoca uma origem comum”.

Primeira fase

Em 1816, a publicação do texto de Franz Bopp sobre a conjugação (*Über das Conjugationssystem der Sanskritsprache, in Vergleichung mit jenem der griechischen, lateinischen persischen und germanischen Sprache*) assinala, tradicionalmente, a data de início da linguística histórica. Isso não por um seu particular valor intrínseco, mas porque a partir dessa obra se movem os primeiros passos da comparação como ciência, sobre a qual a linguística histórica tem o seu fundamento metodológico.

A referência ao valor intrínseco deve-se ao fato de que, realmente, o interesse que move a análise comparativa de Bopp é ainda o de remontar a um estado primitivo, originário, no qual as formas gramaticais se deixam analisar em elementos inexplicáveis em si e não o propriamente histórico, de reconstruir, mediante a comparação das formas documentadas, um precedente estado linguístico comum. Meillet diz que ele ainda é um homem do século XVIII, que pretende remontar à explicação originária dos fatos, dos quais a ciência, por ele

fundada, tornou conscientes os seus sucessores de que é possível apenas compreender o desenvolvimento histórico. De modo que F. Bopp descobriu a gramática comparada procurando uma explicação para os elementos morfológicos indo-europeus, como Colombo descobriu a América procurando o caminho para as Índias. Para dar um exemplo, Bopp, com base em um princípio então difundido, o da constituição de toda oração em três elementos (sujeito, cópula, predicado), explica a origem das desinências como resultado de aglutinações: *dat* = “ele [é] “dante”” (com a cópula subentendida); *potest* = *pot* atributo, *t* sujeito, *es* cópula.

Na sua *Gramática comparativa (Vergleichende Grammatik des Sanskrit, Send, Griechischen, Lateinischen, Litauischen, Altslavischen, Gotischen und Deutschen, 1833)*, retornam os mesmos princípios: a intuição das relações de parentesco é profunda: todas as línguas i.e. ali compõem (exceto, naturalmente, as ainda não descobertas). Mas o interesse quase exclusivo, e não autônomo, pela morfologia, ou melhor, pela flexão, é um grave limite, porque negligencia o estudo das modalidades do desenvolvimento fonético, que constituirá o mais sólido suporte científico da reconstrução do patrimônio comum.

Até aqui, a comparação aparece como um simples instrumento de curiosidade, ou de busca de explicações glotogônicas. Contudo, novos fermentos teóricos amadurecem; assim, a ideia humboldtiana de língua como *enérgeia*, não *érgon*, ou seja, a criação contínua, manifestação do espírito na sua totalidade e não produto de reflexão (ideia precedida pela descoberta de Vico), enquanto as gramáticas de Edmund Hask (1811) e de Jacob Grimm (1812), respectivamente islandesa e germânica, são as primeiras gramáticas históricas assim como a de Bopp é a primeira comparativa. A descoberta do rotacismo germânico, intuída pelo primeiro e formulada pelo segundo, assinala o início do estudo sistemático do desenvolvimento dos fenômenos fonéticos, que dará os seus melhores frutos na segunda metade do século. Mais uma vez, não é a finalidade extralinguística perseguida pelos pesquisadores, nesse caso o espírito nacional na sua pureza, tal como se reflete na língua, mas o meio empregado para esse fim: a coleta precisa de fatos dialetais, nos quais se diferencia o domínio germânico, no seu desenvolvimento das atestações mais antigas, representa a verdadeira contribuição dessas obras para a constituição da ciência linguística.

O momento naturalista

As obras de Bopp, de Hask e de Grimm representam a primeira fase na história dessa ciência, em que se vê de um lado a afirmação da instância comparativa e de outro a formação do interesse pelas modalidades do desenvolvimento fonético, que representam os elementos essenciais para o surgimento do estudo diacrônico das línguas. Todavia, deve-se destacar, quando se fala em linguística histórica como pesquisa sobre a evolução das línguas, que uma diferença profunda separa a linguística do séc. XIX e a do século XX no modo de compreender a natureza dessa evolução.

A noção de «forma orgânica» dos primeiros tempos do romantismo – as línguas como realizações da espiritualidade individual dos povos – passa, próximo ao fim do século, àquela de “organismo natural”, regulado por leis próprias, como todo outro dado da natureza: a gramática é, conseqüentemente, a «doutrina da vida da língua», que se desenvolve segundo

leis análogas às que regulam o mundo físico. O *Compendium der vergleichenden Grammatik der indogermanischen Sprache* (1861), de August Schleicher, constitui a aplicação de tais princípios fundados sobre a concepção hegeliana da história como realização de liberdade consciente, mas essa liberdade não pode ser encontrada na língua, que é por isso objeto de uma sistematização análoga à das ciências naturais (é significativo que Schleicher fosse também botânico e naturalista). Essa obra representa também a primeira organização de uma doutrina linguística indo-europeia, como reconstrução de um estado linguístico não documentado, através da comparação de estados linguísticos historicamente atestados que apresentam correspondências evidentes. Tais correspondências, por ser o resultado de evoluções que podem se percorrer no seu desenvolvimento, permitem que se remonte à uma “língua mãe”, que representa o tronco do qual se ramificam os diversos grupos (*Stammbaum-theorie*: teoria da “árvore-genealógica”). Dessa fase originária, concebida como algo puro e perfeito (note-se a persistência de certas posições iluministas, as quais sobrevivem também, em âmbito teórico, nas concepções de W. Humboldt) move o desenvolvimento diferenciado, visto como «decadência», o qual se verifica, exatamente, segundo determinadas leis naturais.

A idade neogramática

Do conceito de língua como organismo natural, passa-se logicamente ao axioma da validade absoluta das leis fonéticas: é este o axioma que, defendido pela orientação positivista da cultura, domina no último quarto do século XIX. A escola dos neogramáticos (Scherer, Leskien, Osthoff, Brugmann, Delbrück etc) funda a partir dele o entusiasmo com que atua sobre a enorme quantidade de material oferecido pelas diversas filologias, atingindo a sistematização rigorosa do fonetismo indo-europeu que, por sua vez, reforça, com a qualidade dos resultados, a confiança nos seus princípios. Entretanto, diante do aspecto concreto dessa problemática propriamente linguística, adquire-se definitivamente, ainda que não ainda de modo perfeitamente consciente, o sentido da perspectiva histórica. Assim, já em 1875, não aparecem mais em nenhuma publicação, as ingênuas tentativas glotogônicas de tipo boppiano (a redação de fábulas na língua reconstruída), nem se pensa mais no indo-europeu como uma língua perfeita, colocada na origem do desenvolvimento linguístico. Além disso, aparece a exigência de se observar de perto os fatos, ou seja, de se basear não mais na língua escrita, mas sim na observação direta da língua falada (tal já era o valor essencial da gramática lituana de Schleicher). Enfim, a fé na lei comporta o reconhecimento da existência de anomalias e, por mais que se atribua a sua causa a um fator «analógico» entendido também em modo absolutamente mecânico, todavia o fato de mais se admitir, como era frequente em época anterior, a possibilidade de uma mudança «esporádica», ou seja, sem uma causa, passa a dirigir a atenção exatamente sobre essas inegáveis exceções e, portanto, a fazer que se revelasse logo a insuficiência de um princípio tão genérico e impreciso, como é o da “analogia” dos neogramáticos.

Reação ao princípio de “validade absoluta” das leis fonéticas

A passagem da concepção naturalista àquela mais propriamente histórica na consideração da mudança linguística é assinalada, precisamente, pela exigência de explicação das causas das chamadas *anomalias*. Contra a validade absoluta da lei fonética, se erguem as vozes dos empenhados principalmente no estudo de fatos linguísticos românicos (H. Schuchardt, G. Ascoli, e depois J. Gilliéron, o fundador da geografia linguística, precursor de Johannes Schmidt) mais bem documentados e que podiam revelar a quantidade de tais exceções e, definitivamente, aquela variedade do fenômeno língua, para a qual muitas inovações parecem fugir aos rígidos esquemas neogramáticos: a lei linguística não é “natural”, na medida em que opera segundo determinações espaciais e temporais, e não sem uma certa participação da consciência dos falantes. Isso leva à negação da sua “necessidade absoluta”. Nasce os conceitos de interferência, de parentesco secundário ou cultural, ou de afinidade, de etimologia popular, de homofonia e, definitivamente, a exigência de estudar a mudança fonética não isoladamente, mas na palavra que muda. Analisa-se a mudança fonética, portanto, em relação não somente com a influência dos sons vizinhos, mas também com a função semântica da unidade em que se verifica.

Com a nova atenção então voltada aos significados (Geografia linguística, Palavras e Coisas, Onomasiologia), a análise dos fatos linguísticos começa a se tornar efetivamente histórica, no sentido em que, adquirindo consciência da complexidade dos fenômenos e da multiplicidade das causas que concorrem à sua determinação, sente-se a exigência de colocá-lo nas exatas circunstâncias de lugar e de tempo em que foi verificado.

Paralelamente, no plano teórico, o subjetivismo neoidealístico marcava o advento do individual, do subjetivo, do criativo no âmbito da problemática linguística e, portanto, no problema da mudança das línguas. Porém, a arbitrariedade, a aproximação e a incoerência dos resultados são os mais freqüentes frutos de um método dirigido à individuação da “índole” dos povos com base em fatos linguísticos isolados (que mais tarde irá levar a acusação de atomismo à linguística histórica, indevidamente identificada com a linguística idealística, que é apenas uma sua fase) ou, pior ainda, voltados à demonstração de teses extralinguísticas.

A instância estruturalista

Contra um método desse tipo, baseado em uma posição teórica manifestadamente unilateral, como aquele que na língua resultava em afirmar apenas a liberdade e a criatividade subjetiva, deixando de lado completamente o aspecto objetivo que é a condição de tal atividade, o movimento estruturalista se configura como uma reação natural, manifestando, pelo menos na sua fase inicial, um extremismo semelhante, mas em direção oposta.

Uma dialética desse tipo se deve, indubitavelmente, à singular recepção das doutrinas de Ferdinand de Saussure, mesmo se favorecida por certas circunstâncias, como a formulação ainda provisória de tais doutrinas e, por outro lado, a publicação póstuma em forma de um tratado orgânico (o *Cours de linguistique générale*, publicado em 1916, preparado pelos alunos

Bailly e Sechehaye) a qual ainda não tinham atingido, e pela qual teria sido indispensável o esforço de síntese de quem o tinha criado.

Entre os velhos e os novos tempos, cheio de iniciativas inovadoras, mas consciente das conquistas já alcançadas pela linguística e com pleno domínio do método rigoroso dos neogramáticos, F. de Saussure representa a marca que distingue a linguística do século XX. A sua obra representou uma enérgica chamada à objetividade sincrônica da língua, mas está fora de discussão que nela se encontram também as premissas, teóricas e metodológicas, de um método histórico produtivo e autêntico.

Para Saussure, a sistematicidade e o formalismo caracterizam a língua, enquanto objeto de consideração científica (a determinação do objeto específico da ciência linguística é o objetivo constante da sua reflexão, o motivo de fundo da sua problemática): "... partout et toujours ce même équilibre complexe de termes que se conditionnent réciproquement. Autrement dit, la langue est une forme et non une substance"¹¹ (*Cours...*, p. 169). Ou seja, a língua é o patrimônio coletivo de formas fônicas, "significantes", univocamente combinadas com os relativos «significados». Esse patrimônio de signos é organizado em «sistema», na medida em que cada um desses signos deve a sua existência ao fato de entrar em certas relações com os outros. A funcionalidade do sistema, isto é, o que o torna um instrumento apto a funcionar em cada ato de «palavra», é constituída exatamente pelas oposições e correlações intercorrentes entre cada elemento, os quais resultam individuados pelas suas relações diferenciais no conjunto dos elementos similares, mais que pelas suas características positivas: "dans la langue il n'y a que des différences"¹² (*ib.*, 166). Porém, não é menos saussuriana a afirmação da interdependência e, portanto, da substancial indivisibilidade da "palavra" e da "língua", essa considerada o produto e o instrumento daquela, assim como a programação de uma linguística da palavra ao lado de uma da língua.

Na verdade, a enérgica chamada de Saussure à objetividade do sistema, em relação à subjetividade do falante – a quem deu destaque o caráter incompleto da sua obra de sistematização teórica, exatamente em relação à "palavra" – resolveu-se, em conexão, com a polêmica anti-idealista, na definitiva dicotomia sincrônico-diacrônica, operada pelo Estruturalismo.

Nascido, desde o manifesto da escola de Praga, com a precisa e declarada finalidade de uma "linguística sincrônica", esse direcionamento se demonstrou imediatamente empenhado em dar conta das mudanças no próprio ato que criava a noção de equilíbrio estrutural, pronto para se restabelecer automaticamente cada vez que um "fator externo" colocava em crise a harmonia do sistema. Uma posição em que, por "sistema" se entendia racionalistamente – e, portanto, arbitrariamente – se a língua é uma forma da atividade cognoscitiva integral e não só de um seu momento, como racional a abstração das relações diferenciais intercorrentes entre as unidades de uma dada língua. Por elemento "extra-funcional" cada entidade que não pode ser enquadrada na organização simétrica dessas relações, arbitrariamente identificada

¹¹ Tradução: ...em todo lugar e sempre esse mesmo equilíbrio complexo dos termos que se condicionam reciprocamente. Dito de outra forma, a língua é uma forma e não uma substância.

¹² Tradução: na língua há tão somente diferenças.

com a estrutura da língua e por «fator externo» todo elemento não identificável com a exigência, automaticamente ativa no sistema, de distinção das unidades.

Colocadas essas premissas, o Estruturalismo procurou dar também ao problema da inovação uma resposta estrutural, nitidamente distinta das oferecidas pela linguística histórica: segundo o princípio que a «imanência» do método científico requer que se encontre as razões de todos os fenômenos linguísticos, incluindo, pois, a mudança, no interior do sistema linguístico. A insuficiência teórica do princípio metodológico torna-se evidente quando se pensa que o sistema funciona nos atos linguísticos individuais dos falantes, de modo que a palavra, com as suas exigências subjetivas e as suas realizações criativas, representa o impulso contínuo da dinâmica do próprio sistema.

É mérito inegável do Estruturalismo ter definitivamente chamado a atenção para o fato que, para atingir a essência do fenômeno linguístico, é indispensável considerar as unidades relativas também nas mútuas relações que intercorrem entre elas. Mas é também verdadeiro que identificar a exigência imanentista com a abstração do sistema de funções da realidade concreta que a atua, condena a perder de vista o objeto da pesquisa, ou seja, as línguas nas suas fisionomias individuais: a experiência da «glossemática», ou a do Estruturalismo americano, pelo menos nas suas partes mais avançadas, oferecem uma confirmação precisa dessas afirmações.

A linguística histórica hoje

Hoje, a linguística histórica tem condições de dar uma resposta adequada ao porquê da mudança linguística, tornada consciente da complexidade do fenômeno, em relação com a multiplicidade dos fatores e a variedade das suas interações, e por isso capaz de utilizar a contribuição dada por diversas orientações, equilibrando as instâncias unilaterais de cada uma delas. A primeira condição para isso é não exasperar a oposição língua-palavra (e muito menos a de diacronia-sincronia). Quando, a propósito da mudança da estrutura sintética àquela analítica, ocorrida no desenvolvimento de quase todas as línguas indo-europeias, Meillet sustentava a intervenção de “fatores afetivos” de um lado e «lógicos» de outro, acrescentando explicitamente que se esses se afirmam é porque a inovação é imediatamente assumida por indivíduos que participam das mesmas condições históricas, o que significava unir, sobre o plano histórico concreto¹³ a presumida cisão entre língua e palavra; significava que o aluno de Saussure tinha compreendido a lição de quem, mesmo tendo criado as premissas da “*langue en elle même et par elle même*”¹⁴ – e é interessante relevar que hoje temos como certo que a formulação do princípio de imanência nesses termos não é saussuriano – afirmava, todavia, a interdependência entre língua e palavra, definindo a língua tanto como um instrumento quanto como um produto da palavra (*Cours...*, p. 37).

Eliminada tal oposição, é necessário não criar uma nova entre fator “interno” e «externo» de mudança, compreendendo a qualidade de “fator externo” à maneira proto-

¹³ Mesmo se mediante o trâmite sociológico que em plena idade estruturalista reaparece, não por acaso, no conceito de “norma” de Coseriu.

¹⁴ Tradução: língua nela mesma e por ela mesma.

estruturalista, ou seja, como equivalentes de linguístico e extra-linguístico. Quando o falante, com suas reações subjetivas, ou o substrato, com as suas realizações, são considerados como elementos estranhos à língua, entendida como sistema abstrato de relações diferenciais, é evidente que para nós se fecha qualquer possibilidade de compreender a dinâmica do próprio sistema porque na realidade os fatores “internos”, como as exigências dinâmicas do sistema, apresentam-se ao falante tanto quanto os “externos”, como as influências do substrato, adstrato e superestrato. Se uma língua pode determinar “reações” em uma outra, com a qual mantenha contato ou que a tenha suplantada, isso acontece porque o patrimônio espiritual, que aquela língua exprime, assume junto aos falantes da outra comunidade um certo prestígio, de ordem política ou cultural. A mudança que deriva disso é de natureza externa, porque o empréstimo, ou a reação, não se produz mecanicamente no sistema, mas documentam um momento ativo no comportamento linguístico daquela comunidade. Que o fator endógeno de mudança seja atribuído ao falante é ainda mais evidente: se ele responde à exigência, inerente ao sistema, de reforçar e enriquecer as possibilidades de expressão. É claro, por outro lado, que tal reforço ou enriquecimento tem origem e se realiza no plano do ato linguístico e, precisamente por obra dos elementos extra-funcionais, que a atividade do falante continuamente introduz no sistema. Assim, no léxico como na morfologia, momentos estilísticos, assumidos e tornados comuns pela comunidade, tornam-se fatos da língua. Analogamente, no plano fonético, tom e tempo do discurso (quando já não sejam elementos “funcionais” no sistema) e a alusividade (a “fonética impressiva” de M. Grammont) são a causa de “variantes”, prontas a adquirir relevância no plano fonemático.

Uma das tarefas mais importantes assumida pela “crítica semântica” de A. Pagliaro é, exatamente, a de indagar a língua como função, que põe em ação a funcionalidade do sistema, para procurar nos fenômenos linguísticos, em que tenham assumido forma momentos subjetivos excepcionais, as causas e os modos da aceitação do fato estilístico como elemento funcional.

A regularidade da mudança, uma vez que ela tenha se afirmado, é precisamente a expressão da sua correlação com os elementos funcionais da mesma ordem em um novo equilíbrio; mas por que esse tenha podido se afirmar deve ser procurado na comunhão dos falantes, quer dizer, na sua comum história linguística.

EXERCÍCIO

Esquematize as diversas etapas da linguística histórico-comparativa.

Guia de Leitura Texto 3: O Método Histórico-Comparativo

- (1) Qual fato pode ser considerado o mais determinante para a fundação da Linguística Comparada?
- (2) Qual a principal obra responsável pela difusão do conhecimento do sânscrito na Europa? Quem foi seu autor?
- (3) Por que a aproximação dos europeus desta língua causou tanto impacto nos estudos comparativos? Cite, ao menos, três fatores que expliquem esse impacto.
- (4) Qual a explicação, dada na Antiguidade, para as semelhanças e proximidades gramaticais entre o latim e o grego?
- (5) Quando os estudos comparativos avançaram para além da comparação superficial entre línguas, encetada na Antiguidade?
- (6) Em que ocasião o estudo da linguística indo-européia começou a estruturar-se nos moldes de uma disciplina moderna? Qual hipótese impeliu os acadêmicos nesta nova direção de estudo?
- (7) Quais os méritos da declaração de Jones sobre o sânscrito?
- (8) Em que consistiram os estudos do linguista dinamarquês Rasmus Rask?
- (9) Qual a contribuição ou a influência do estudo de Franz Bopp, publicado em 1816, sobre o campo dos estudos comparativos?
- (10) Qual a diferença da *Deutsche Grammatik*, de Jacob Grimm, em relação às obras de seus antecessores?
- (11) Explique, sinteticamente, o funcionamento da chamada *Lei de Grimm*.
- (12) Qual o marco de criação da Filologia Românica?
- (13) Qual a relevância do *Satiricon* de Petrônio para a compreensão da evolução histórica da língua latina?

TEXTO 4

A RECONSTRUÇÃO DO INDO-EUROPEU

Jacyntho Lins Brandão

Vamos retomar a família indo-europeia para seguir alguns dos passos a partir dos quais se reconstitui uma protolíngua. Há três razões principais para essa escolha: (a) como você já sabe, trata-se de uma família bastante estudada desde o século XIX, de cuja protolíngua não temos nenhum registro escrito, o que faz com que seu conhecimento dependa inteiramente da aplicação do método comparativo; (b) é nessa família que se encontra o português, o que ajuda a compreender, em termos diacrônicos, algumas de suas categorias gramaticais; (c) da comparação do proto-indo-europeu com as línguas indo-europeias percebe-se como a mudança linguística se processa sem pressupostos teleológicos.

A reconstituição de uma protolíngua não documentada tem como primeiro requisito a possibilidade de, comparativamente, a partir de um conjunto de línguas que se pretenda dela sejam derivadas, estabelecer o que define uma língua, a saber: (a) um léxico; (b) um sistema fonológico; (c) um sistema de morfológico; (d) padrões sintáticos. Atenção: como as línguas encontram-se em processo constante de variação e mudança, não se trata de estabelecer essas categorias para todo o conjunto, mas de deduzir do conjunto quais seriam os traços que se encontrariam na protolíngua reconstituída, capazes de explicar os resultados observáveis nas diferentes línguas dela procedentes.

Um segundo requisito é que a protolíngua deve ser tipologicamente viável. Como veremos adiante, existem alguns padrões linguísticos (o que se denomina “tipos linguísticos”), deduzidos da observação e comparação das línguas existentes. Por exemplo, recordando algo a que já se fez referência: nem todas as línguas apresentam artigos, mas, se uma língua tem apenas um tipo de artigo, este será o definido. Noutros termos, de uma perspectiva diacrônica: tudo parece indicar que, quando se criam artigos, primeiro se cria o definido e apenas depois, se for o caso, o indefinido. Nesse sentido, caso se reconstitua uma língua que possua apenas artigos indefinidos isso representará, em princípio, uma dificuldade tipológica que põe em suspeição a própria reconstituição.

Na sequência, vamos abordar alguns dos passos da reconstituição do indo-europeu, enfatizando o que diz respeito às principais categorias morfológicas nominais e verbais, pois,

como você verá, seu conhecimento ajuda a entender as categorias de número, gênero, caso, aspecto, tempo e modo nas línguas indo-europeias modernas, incluindo o português.

1 Como se reconstroi uma língua

O primeiro passo na reconstituição de uma língua a partir do método comparativo é bastante simples: a comparação de palavras que, cobrindo uma mesma esfera semântica, mantenham entre si alguma semelhança, de modo que se possam propor processos de mudança fonética que pareçam plausíveis.

Vamos experimentar seguir um pouco dos passos desse processo, observando os termos registrados no quadro abaixo e agrupando-os em vista de sua maior ou menor semelhança:¹⁵

QUADRO 1
Quadro comparativo

	I	II	III	IV	V	VI	VII	VIII	IX
1	cem	cent	ciento	hundert	centum	sută	cento	yüz	cent
2	oito	huit	ocho	acht	octo	opt	otto	sekiz	vuit
3	noite	nuit	noche	Nacht	nocte(m)	noapte	notte	gece	nit
4	filha	fille	hija	Tochter	filia	fiică	figlia	kiz	filla
5	folha	feuille	hoja	Blatt	folia	foaie	foglia	yaprak	fulla
6	figo	figue	higo	Feige	ficu(m)	smochin	fico	incer	figa
7	saber	savoir	saber	wissen	sapëre	şti	sapere	bilmek	saber
8	vida	vie	vida	Leben	uita	viață	vita	hayat	vida
9	orelha	oreille	oreja	Ohr	auricula	ureche	orecchia	kulak	orella
10	velha	vieille	vieja	alte	uetŭla	veche	vecchia	yaşli	vella

português

Agora siga passo a passo:

(a) Com certeza, você descobriu uma das colunas que não apresenta nenhuma semelhança com as demais (repare, por exemplo, as palavras para *oito* e *noite*), nela estando arroladas as palavras do turco (anote no quadro onde ele se encontra).

(b) Há outra coluna que, apesar da semelhança dessas palavras (*oito* e *noite*) com as das demais colunas, nos outros itens também apresenta diferenças notáveis: é o alemão (registre também isso no quadro). As demais línguas são todas procedentes do latim,

¹⁵ Algumas observações sobre algumas convenções ortográficas: 1. na coluna VI, a letra ð grafa o fonema /ə/, semelhante ao æ do inglês *father*, ð' representa o fonema /ʃ/, equivalente ao 'ch' do português, e 't' o fonema /ts/; na coluna VIII, ðø representa /ʃ/ e ðg' não grafa um fonema específico, apenas alongando a vogal que o precede.

cujos termos se encontram registrados na coluna V (as línguas dele derivadas, no quadro, são, pela ordem: português, francês, espanhol, romeno, italiano e catalão).

(c) Na linha 1 será fácil você perceber as línguas que modificaram mais a palavra latina. Agora observe o seguinte: a letra 'c' em latim clássico representa a velar desvozeada, ou seja, /k/. Assim, para chegar às formas românicas, a primeira mudança já se produziu no próprio latim vulgar (o latim falado donde procedem as línguas românicas), no seguinte sentido: /ke/ > /se/. Tendo isso em vista, a forma do romeno, que, em princípio, parece mais distanciada, mostra o quanto tem de foneticamente semelhante com as demais, ainda que grafie a inicial com 's'. Organize as línguas na sequência das que são mais ou menos semelhantes ao latim 'centum', observando também o vocalismo.

(d) Agora observe nas linhas 2 e 3 a evolução da sequência latina *-ct-* nas diferentes línguas, nas palavras *octo* e *nocte(m)*, e descreva-a abaixo:

QUADRO 2

Evolução da sequência *-ct-* nas línguas românicas

<i>-ct- > -it-</i>	
<i>-ct- > -ch-</i>	
<i>-ct- > -tt-</i>	
<i>-ct- > -pt-</i>	

e) Preste atenção agora com o que acontece com o 'f' inicial latino e nas linhas 4, 5 e 6: ele se conserva em todas as línguas, com exceção de uma. Qual? Trata-se de uma mudança que se debita ao substrato (ou seja, à língua falada numa determinada região antes da implantação de outra): neste caso, o substrato basco, que não só não é uma língua latina, mas nem mesmo indo-europeia. A substituição do /f/ inicial por uma aspiração (/h/) registra-se lentamente a partir do século IX, na zona ao norte de Burgos, e só se impõe literariamente no século XV, diferenciando, assim, o espanhol de todas as outras línguas românicas (ELIA, 1979, p. 94-96).

(f) Ainda nas linhas 4 e 5, verifique a evolução da sequência latina *-li-*vogal: considerando que os dígrafos *-lh-* (português), *-ll-* (catalão) e *-gl-* (italiano) representam o mesmo fonema /λ/, que línguas mais se afastaram do latim?

(g) Nas linhas 6, 7 e 8, observe a evolução das consoantes desvozeadas intervocálicas, a saber, /k/, /p/ e /t/ e você constatará que as línguas se organizam de um modo bastante regular: as que mantêm a consoante latina; as que a mudam para a correspondente vozeada; uma língua que muda a vozeada para a correspondente fricativa (/p/ > /b/ > /v/) ou elimina completamente a consoante (o romeno, nas linhas 6 e 7, adotou outros

termos, por isso não é representativo das mudanças fonéticas). Feita essas observações, você constatará que as línguas da Península Ibérica vozeiam regularmente as desvozeadas (o fenômeno chamado de “sonorização das surdas intervocálicas”, que costuma ser atribuído ao substrato celta), sendo provável que o francês também tenha conhecido esse estágio, antes de chegar às formas atuais (podendo-se, portanto, dizer que se trata de tendência característica da România ocidental); por outro lado, o italiano e o romeno tendem a manter as desvozeadas intervocálicas do latim (o que seria um fenômeno típico da România oriental).

(h) Enfim, nas linhas 9 e 10 você encontra dois fenômenos semelhantes: na linha 9, a palavra latina para ‘orelha’, *auris*, é substituída por seu diminutivo, *auricŭla*, que, por ser proparoxítona, muda para **auricla*. No *Apendix Probi*, uma lista de formas incorretas do latim feita por um autor anônimo provavelmente do século III d.C., temos um importante testemunho sobre a forma dessa palavra no latim da época, já que ele anota “*auris non oricla*”. Do mesmo modo, o diminutivo de *uetus*, ‘velho’, é *uetŭlus*, sendo este último que dá origem a **vetlus* donde procede *veclus*, outra palavra registrada no *Apendix Probi*: “*vetulus non veclus*.” Portanto, o ponto de partida das palavras apresentadas nas linhas 9 e 10 encontra-se nas formas vulgares *oricla* e *vecla*. Agora observe o tratamento do encontro consonantal *-cl-* e veja como as línguas se distribuem entre as duas partes da România já referidas, a ocidental e a oriental (o espanhol apresenta mais uma mudança própria: *-ll-* > *-j-*).

No caso das línguas românicas, temos a situação ideal de contarmos com registros da língua de origem, ao lado das derivadas. Isso, todavia não seria suficiente, se não se pudessem entender os fenômenos de mudança fonética, que, saliente-se, não são uma camisa-de-força, pois interferem no processo também fatores de ordem pragmática e cultural.

2 A reconstituição do indo-europeu

Evidentemente, quando a tarefa é reconstituir uma língua de que não se tem nenhuma documentação e que teria sido falada, provavelmente, há sete mil anos, tudo se torna mais difícil, embora não seja impossível.

No quadro abaixo, vamos fazer o mesmo exercício de comparação:

QUADRO 3
Quadro comparativo

	1	2	3	4	5	6	7	8
	noite	dois	três	cinco	pé	coração	que/quem*	português
I	noctis	duo	tres	quinque	pedis	cordis	quid	
II	nuktós	dúo	treís	pénta	podós	kardía	tí	
III	night	two	three	five	foot	heart	what	
IV	oíche	dhá	trí	cúig	cos	chroí	cad	
V	gaua	bi	hiru	bost	oinez	bihotza	zer	
VI	notsh	dva	tri	piat	fut	sierdtsie	kto*	
VII	nakti	dvau	trayah	pañca	pad	hṛdaya	kás*	
VIII	éjszaka	ket	három	öt	láb	szív	mi	
IX	naktis	du	trys	penki	péda	širdis	kas*	
X	natt	tva	tre	fem	fot	hjärta	vad	

Faça assim:

- Comece pelas colunas 1, 2 e 3, com as palavras para 'noite', 'dois' e 'três' e verifique quais as duas línguas que não apresentam semelhanças com as demais (a primeira é o basco, a outra, o húngaro, nesta ordem – anote na última coluna). Agora confira se a mesma tendência se repete nas outras colunas (se necessário, corrija sua opção anterior).
- Agora observe que, na coluna 2, a maior parte das palavras para 'dois' começam com a dental vozeada /d/, havendo apenas duas que apresentam a correspondente desvozeada /t/; nas colunas 3 e 4, onde a maior parte tem a labial desvozeada /p/, nas mesmas duas encontramos /f/; na coluna 6, as mesmas duas línguas têm a aspirada /h/ onde outras trazem a velar desvozeada /k/ ou a sibilante /s/. Fazendo isso, você identificou a chamada "rotação germânica" e as duas línguas germânicas do quadro, a saber, o inglês e o sueco (escreva os dois nomes).
- Na mesma coluna, observe agora a distribuição entre /k/ e /s/, as que apresentam a última solução sendo, pela ordem, o russo e o lituano (escreva os nomes).
- Na última coluna, você tem um exemplo de outra distribuição: onde a língua I tem uma labiovelar /k^w/, as demais línguas apresentam três alternativas: /k/, /w/ ou /t/. Compare agora com o tratamento da última sílaba da coluna IV e você descobrirá que as línguas 1 e 2 mantiveram a correlação /k^w/ ≈ /t/, sendo a primeira o latim e a segunda o grego (não se esqueça de anotar os nomes).

(e) Para completar as informações, agora anote as demais línguas do quadro, que são, pela ordem, o irlandês e o sânscrito.

É assim que se trabalha comparativamente, caminhando passo a passo e estabelecendo tendências. A partir disso é que se fazem as propostas de reconstituição do sistema fonológico do indo-europeu e dos diferentes ramos de sua família.

Todavia, uma vez verificada a semelhança no nível meramente lexical, é preciso avançar pelo estudo das categorias gramaticais, pois é isso que dá consistência ao modelo. Vamos então prosseguir.

2.1 A categoria de número

Em seu conjunto, registram-se nas línguas indo-europeias três números: singular, plural e dual. Como você vê, a oposição singular/plural não é a única, havendo línguas que, além do dual, apresentam ainda outros números, o que exploraremos mais à frente. Tipologicamente, entre os três que agora nos interessam, há uma organização hierárquica: (a) se uma língua possui o plural, possuirá também o singular; (b) se possui o dual, possuirá também o plural e o singular.

É importante ressaltar que todas as línguas têm formas de indicar o número, mas considera-se que possuem essa categoria gramatical apenas aquelas que atendem a dois requisitos: (a) apresentam alguma forma de marcação do número através de afixos, ou seja, um procedimento de natureza morfológica; (b) têm o número como uma categoria de concordância, o que significa dizer que ela exerce uma função sintática. No grupo indo-europeu, ambas as exigências se cumprem, como você pode constatar dos exemplos abaixo, em que tanto o substantivo apresenta sufixos de marcação de número, quanto seus determinantes e o verbo com ele concordam:

QUADRO 4

Exemplos de singular/plural/dual enquanto categorias gramaticais

	Singular	Plural	Dual
Português	esta fruta está madura.	estas frutas estão maduras.	x
Francês	<i>ce fruit est mûr.</i>	<i>ces fruits sont mûrs.</i>	x
Inglês	<i>this fruit is ripe.</i>	<i>these fruits are ripe.</i>	x
Sânscrito	<i>etat phalam pakvam.</i>	<i>etāni phalāni pakvāni.</i>	<i>ete phale pakve.</i>
Grego	<i>ho karpòs hoûtos hóriós esti.</i>	<i>hoi karpòi hoûtoi hóriói eisi.</i>	<i>tō karpō toútō horiō estón.</i>

Observe:

- (a) Nos três primeiros exemplos (português, francês e inglês) há dois números (singular/plural), enquanto nos dois últimos (sânscrito/grego) encontramos três (singular/plural/dual);
- (b) Em português e francês tanto o demonstrativo (*esta/ce* x *estas/ces*), quanto o verbo (*está/est* x *estão/sont*) e o predicativo (*madura/mûr* x *maduras/mûrs*) concordam em número com o sujeito (*fruta/fruit* x *frutas/fruits*);
- (c) No inglês, o demonstrativo (*this/these*) e o verbo (*is/are*) também apresentam concordância de número com o sujeito (*fruit/fruits*), mas não o predicativo (*ripe*), já que o adjetivo nesta língua é invariável;
- (d) O sânscrito dispensa o verbo de ligação (*etat phalam pakvam* = esta fruta madura), o demonstrativo (*etat/etāni/ete*, ‘este/estes/estes dois’) e o predicativo (*pakvam/pakvāni/pakve*) concordando em número com o sujeito (*phalam/phalāni/phale*, ‘fruto/frutos/dois frutos’).
- (e) Em grego, todos os elementos da oração realizam a concordância de número, a saber: o artigo (*ho/hoi/tō*, ‘o/os/os dois’), o demonstrativo (*hoûtos/hoûto/toútō*, ‘este, estes, estes dois’), o predicativo (*hórios/hórioi/horiō*) e o verbo (*estí/eisí/estón*, ‘ele é/eles são/eles dois são’).

Naturalmente, o que se diz no dual em sânscrito e grego pode-se dizer também em português, francês e inglês com o acréscimo do numeral ‘dois’ (*estas duas frutas estão maduras/ces deux fruits sont mûrs/these two fruits are ripe*), mas então não se trata de número gramatical, a concordância fazendo-se no plural, uma vez que não existe a oposição entre plural e dual. Acrescente-se que em grego sempre se pode usar o plural no lugar do dual, tendo este último número um uso bastante refeito, de modo que *tō karpō toútō horiō estón* (dual) equivale a *hoi dúo karpōi hoûtoi hórioi eisi* (plural, sendo *dúo* o numeral ‘dois’). No grego comum (*koiné*), já no fim da Antiguidade, bem como no grego moderno, o dual não existe mais.

Nas línguas indo-europeias, o dual apresenta-se em algumas das línguas de que dispomos de registros que remontam à Antiguidade – como, além do sânscrito e do grego, também no avéstico, no gótico e no antigo eslavo eclesiástico –, não se registrando, contudo, em outras igualmente antigas, como o latim e o hitita. Manteve-se ainda nos registros mais arcaicos do islandês e conserva-se, ainda hoje, no esloveno (*lipa/lipe/lipi*, ‘tília/tílias/duas tílias’). O fato de que se encontre atestado em diversos grupos (indo-iraniano, grego, germânico e eslavo) leva a supor que tenha sido gramaticalizado em fase bastante remota do indo-europeu, embora posterior à separação do grupo anatólio, uma vez que nas línguas desse grupo não há qualquer traço do dual. Ressalte-se que em latim encontramos reminiscência do dual no termo *ambō*, ‘ambos’.

Considerando os fatos referidos, podemos representar a categoria de número no proto-indo-europeu da seguinte forma:

QUADRO 5
A categoria de número no proto-indo-europeu

Singular (apenas um)	Plural (mais de um)	
	Dual (pares)	Plural (mais de dois)

Em geral, nas línguas que o apresentam, o dual, enquanto um desdobramento do plural, é uma oposição que se aplica a um tipo determinado de seres ou coisas, as que aparecem aos pares, como irmãos gêmeos, um casal, os olhos, as orelhas etc. (em português, alguns termos que têm forma de plural, como ‘óculos’, guardam alguma ideia de dual, tanto que a concordância é flutuante, podendo-se dizer tanto ‘meus óculos quebraram’, considerando-se que se trata de um objeto composto de um óculo esquerdo e um óculo direito, quanto ‘meu óculos quebrou’, tendo em vista que se trata de um par de óculos). É esse componente semântico que dá ao dual uma identidade mais específica, deduzida de certa visão de mundo: há o que se mostra sempre como unidade (por exemplo, o céu, o mar, a terra), ou se apresenta como tal em determinadas situações (um homem, um animal, uma árvore); há o que se apresenta sempre aos pares (olhos, pernas, braços), ou, eventualmente, assim se mostra em determinadas situações (dois homens, dois animais, duas árvores); finalmente, há tudo o que ultrapassa a unidade e/ou o par.

Outra forma de indicar o número presente nas línguas indo-europeias e provavelmente no próprio indo-europeu é o coletivo. Neste caso, trata-se de uma palavra que expressa uma pluralidade de indivíduos da mesma espécie, considerados todavia de uma perspectiva de conjunto: ‘rebanho’, ‘cardume’, ‘cacho’, ‘enxame’, ‘tropa’, ‘multidão’. Além da existência de itens lexicais com a significação própria de coletivo (podendo, inclusive, em alguns casos, flexionar-se no plural: ‘rebanhos’, ‘tropas’), parece ter havido um sufixo de coletivo no indo-europeu (-aH), de onde proveio a desinência de plural do neutro (cf. latim *templum/templa*, ‘templo/templos’; grego *biblíon/bibλία*, ‘livro/livros’). A memória de que se tratava não de um plural comum, mas de um coletivo, conservou-se em grego antigo, já que o sujeito no neutro plural concorda com o verbo no singular: *tò téknon trékhei*, ‘o menino corre’/*tà tékna trékhei*, ‘os meninos correm’, com o verbo *trékhein*, ‘correr’, mantendo-se sempre na terceira pessoa do singular (o sentido da construção de plural sendo algo equivalente ‘a meninada corre’). O mesmo sentido de coletivo foi preservado em *bibλία*, ‘livros’, desde quando o empréstimo

grego a um sem número de outras línguas passou a designar 'a Bíblia' enquanto o conjunto de livros sagrados de judeus e cristãos.

2.2 A categoria de gênero

Em termos gerais, a categoria de gênero pode ser considerada gramatical quando, além da eventual existência de afixos que o expressam (no plano morfológico), exerce ele um papel na concordância nominal ou verbal (função sintática). Seria mais correto, portanto, considerá-la como uma "classe de concordância", a exemplo do que você já viu com relação às línguas nigero-congolesas, dentre as quais o quimbundo. Como nas línguas indo-europeias o verbo concorda com o sujeito apenas em número e pessoa, o gênero gramatical tem sua função restrita à concordância nominal.

Nas línguas indo-europeias historicamente atestadas, encontramos cinco situações:

1. Línguas que não conhecem nenhuma distinção de gênero, como o persa moderno, o tadjique, o assamês (todas do grupo iraniano), o bengali (grupo índico) e o armênio (língua isolada);
2. Línguas que fazem duas distinções de gênero, podendo ser eles:
 - 2.1. comum e neutro, como no hitita (grupo anatólio), no dinamarquês, no holandês, e no sueco (todas estas do grupo germânico);
 - 2.2. animado e inanimado, que é o caso do bretão (grupo celta);
 - 2.3. masculino e feminino, como no letão (grupo báltico), no hindi, no români (grupo índico), no pachto (grupo iraniano) e na maior parte das línguas românicas (grupo itálico), a saber, galego, português, espanhol, catalão, francês e italiano;
3. Línguas que admitem três distinções de gênero – masculino, feminino e neutro –, a exemplo do avéstico (grupo iraniano), do sânscrito (grupo índico), do búlgaro (grupo eslavo), do islandês, do inglês, do alemão (grupo germânico), do latim, do romeno (grupo itálico) e do grego (língua isolada);
4. Línguas que admitem quatro distinções de gênero – a saber, masculino animado, masculino inanimado, feminino e neutro –, como o russo, o tcheco, o eslovaco, o croata, o sérvio, o bósnio e o montenegrino (todas do grupo eslavo);
5. Uma língua, o polonês (grupo eslavo), com cinco distinções de gênero – masculino pessoal, masculino animado não-pessoal, masculino inanimado, feminino e neutro.

A partir dessa relação pode-se concluir: (a) não existe uma distribuição coerente entre os diferentes critérios e os grupos linguísticos, a não ser com relação aos itens 4 e 5, já que se trata de desdobramentos próprios das línguas eslavas (mas observe-se que o búlgaro se encontra no item 3); (b) há duas ordens de motivação semântica envolvidas nas diferentes combinações, a saber, a oposição animado/inanimado e a oposição masculino/feminino. Parece que é essa existência de duas ordens de critérios de base diversos que torna a categoria de gênero um fato complexo no grupo indo-europeu, o que se poderia representar assim:

QUADRO 6

Combinações dos critérios animado/inanimado e masculino/feminino no gênero gramatical das línguas indo-europeias

Animado (ou Comum)			Inanimado (ou Neutro)
Masculino		Feminino	
Animado			
Pessoal	Não-pessoal		

Em termos diacrônicos, parece razoável admitir que a ordem de criação de gêneros teria se processado no seguinte sentido:

- (a) num primeiro estágio, a oposição seria entre animado e inanimado, o que corresponde à situação verificada no hitita;
- (b) num segundo momento, no interior do animado, se teria processado à divisão entre masculino e feminino, como se encontra em sânscrito, grego, latim, gótico e proto-eslavo, gerando um esquema tripartido (masculino x feminino x neutro);
- (c) num terceiro estágio, próprio das línguas eslavas, procede-se a uma nova divisão no interior do masculino, separando as palavras em animados e inanimados (masculino animado x masculino inanimado x feminino x neutro);
- (d) num quarto momento, uma língua eslava, o polonês, efetuou uma nova separação no interior do masculino animado, dividindo-o em pessoal (humanos machos) e não pessoal (animais machos).

Seria necessário, também diacronicamente, percorrer, no sentido inverso, a redução das oposições de gênero em algumas línguas:

- (a) com exceção do hitita, parece razoável reconstituir para as protolínguas dos demais ramos a existência de três gêneros: masculino, feminino e neutro (o item b acima);
- (b) algumas línguas (como o sueco e o bretão) abandonaram a distinção masculino/feminino, retornando a uma oposição binária (animado ou comum/inanimado ou neutro);
- (c) outras línguas (como maior parte das românicas, o hindi e o pachto), conservaram apenas a distinção masculino/feminino;

(d) finalmente, há línguas que eliminaram completamente a distinção de gêneros (persa, bengali, armênio).

A existência do modelo tripartido (masculino/feminino/neutro) no avéstico, no sânscrito, no grego, no latim, no proto-eslavo e no protogermânico sugere que o indo-europeu já conhecia esse modelo antes da separação desses grupos, mas após a constituição do anatólio como grupo à parte (em que prevalece a divisão gênero comum/gênero neutro). A classe dos adjetivos, os determinantes por excelência, é esclarecedora quanto a isso. Em grego e em latim, por exemplo, há dois tipos de adjetivos: os que apresentam duas formas (com o masculino/feminino opondo-se ao neutro) e os que têm três formas (masculino x feminino x neutro):

QUADRO 7
O gênero dos adjetivos em grego e latim

Gênero	Grego			Latim		
	masculino	feminino	neutro	masculino	feminino	neutro
Triformes	<i>agathós</i> , 'bom'	<i>agathē</i>	<i>agathón</i>	<i>bonus</i> , 'bom'	<i>bona</i>	<i>bonum</i>
Biformes	<i>alethēs</i> , 'verdadeiro'		<i>alethēs</i>	<i>brevis</i> , 'breve'		<i>breve</i>

Observe-se que os adjetivos biformes mantêm um modelo mais arcaico, anterior à separação do animado em masculino e feminino, o neutro correspondendo ao inanimado, enquanto os triformes já criaram uma forma própria do feminino na antiga classe do animado. No que diz respeito ao latim, como as línguas românicas, à exceção do romeno, não conservaram o gênero neutro, gerou-se um esquema em que os adjetivos podem ser biformes (masculino/feminino) ou uniformes (sem distinção de gênero), estes últimos dando continuidade à classe dos adjetivos biformes do latim:

QUADRO 8
O gênero dos adjetivos em português, espanhol, francês e italiano

Gênero	Português		Espanhol		Italiano	
	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.
Biformes	bom	boa	bueno	buena	buono	buona
Uniformes	breve		breve		breve	

Os adjetivos uniformes representam, como se vê, o ponto de chegada da antiga divisão entre animados e inanimados, ou seja, neles se eliminou a distinção do gênero enquanto classe de concordância, uma vez que essas línguas não mantiveram o neutro: 'um discurso breve/uma vogal breve'. Em alguns casos, essa distinção é recuperada, como aconteceu no francês, em que, a partir do mesmo termo latino, se retomou a distinção entre masculino (*bref*) e feminino (*breve*): *un discours bref/une voyelle brève*.

Vamos buscar entender a motivação semântica para os dois critérios básicos de distinção de gênero no indo-europeu. O mais recente é fácil de compreender: o masculino designa os seres animados machos; o feminino, os seres animados fêmeas. Quanto à distinção mais antiga, animado x inanimado, alguns esclarecimentos são necessários.

Antes de tudo, é preciso considerar que a categoria gramatical de gênero visa a organizar o léxico de uma língua tanto em termos semânticos quanto funcionais. Do ponto de vista semântico, a organização lexical reflete a organização do mundo tal qual percebida por uma determinada cultura, ou seja, tal qual representada em seu imaginário. Nesse sentido, no conjunto dos seres, podemos considerar que o reconhecimento de uma parte deles como [+ animado], em oposição a outra parte tida como [- animado], é que teria dado origem à primeira grande divisão do léxico indo-europeu em duas partes, compreendendo no gênero animado os seres vivos, que se movem e se reproduzem, o que se poderia representar assim:

QUADRO 9
Critérios semânticos da classificação em animado/inanimado

	Animado	Inanimado
Humanos	+	-
Animais	+	-
Árvores	+	-
Astros	+	-
Filhotes	-	+
Frutos	-	+
Objetos	-	+

Observe que é a oposição produtor/produto que determina que humanos, animais e árvores sejam [+ animado], enquanto filhotes e frutos seriam [- animado]. Em algumas línguas atestadas encontramos alguma reminiscência remota dessa distinção primeira, em exemplos como:

QUADRO 10
Distribuição entre masculinos/femininos e neutros que remontam à distinção animado/inanimado

Termos	Língua	Animados		Inanimados	Critério predominante
		Masculino	Feminino	Neutro	
Pai/mãe/bebê	Grego	<i>patēr</i>	<i>mētēr</i>	<i>téknon</i>	Progenitores x gerado
	Alemão	<i>Vater</i>	<i>Mutter</i>	<i>Kind</i>	
Figueira/figo	Grego		<i>sukē</i>	<i>sûkon</i>	Árvore x fruto
Pereira/pera	Latim	<i>pirus</i>		<i>pirum</i>	
Macieira/maçã	Latim	<i>malus</i>		<i>malum</i>	
	Russo		<i>iablonia</i>	<i>iabloko</i>	
	Grego	<i>mēlīs</i>		<i>mēlon</i>	

Poeta/poetisa/ poema	Grego	<i>poiētēs</i>	<i>poiētris</i>	<i>poiēma</i>	Produtor x produto
-------------------------	-------	----------------	-----------------	---------------	-----------------------

Fontes: Meillet, 1948, p. 211-229; Gramkrelidze e Ivanov, 1995, p. 218-219.

Ilustrativo também é observar, em línguas que apresentam a tripartição de gêneros, como se distribuem alguns nomes de elementos da natureza, o que parece guardar alguma motivação de ordem mais arcaica:

QUADRO 11
Gêneros de elementos da natureza em latim, grego e alemão

Elemento da natureza	Línguas	Animados		Inanimados
		Masculino	Feminino	Neutro
Ceu	Latim	<i>Caelus</i>		<i>caelum</i>
	Grego	<i>ouranós</i>		
	Alemão	<i>Himmel</i>		
Sol	Latim	<i>sol</i>		
	Grego	<i>hélios</i>		
	Alemão		<i>Sonne</i>	
Lua	Latim		<i>luna</i>	
	Grego		<i>selēnē</i>	
	Alemão	<i>Mond</i>		
Estrela	Latim		<i>stella</i>	
	Grego	<i>astēr</i>		
	Alemão	<i>Stern</i>		
Terra	Latim		<i>terra</i>	
	Grego		<i>gē</i>	
	Alemão		<i>Erde</i>	
Fogo	Latim	<i>ignis</i>		
	Grego			<i>pûr</i>
	Alemão			<i>Feuer</i>
Água	Latim		<i>aqua</i>	
	Grego			<i>húdor</i>
	Alemão			<i>Wasser</i>
Mar	Latim			<i>mare</i>
	Grego	<i>póntos</i>	<i>thálassa</i>	<i>pélagos</i>
	Alemão		<i>See</i>	
Rio	Latim			<i>flumen</i>
	Grego	<i>potamós</i>		
	Alemão	<i>Fluß</i>		

Repare como até o elemento 'terra' não há, com a exceção do latim, ocorrência de palavras do gênero neutro. O caso do latim é emblemático: *caelum*, neutro, designa o céu enquanto a calota sobre a terra, enquanto *Caelus*, masculino, o Céu personificado como um

deus. Essa seria uma explicação para a concentração dos nomes relativos ao céu, aos corpos celestes e à terra no masculino ou feminino, uma vez que eles são considerados entidades vivas, logo, do gênero animado. Quando se trata dos elementos naturais terrestres – fogo, água, mar e rio – o número de ocorrências no neutro cresce, havendo mesmo a possibilidade, como no caso das palavras que em grego nomeiam o mar, de contar-se com termos dos três gêneros. Todos esses exemplos ilustram como a sucessão dos dois critérios de distinção de gênero fez com que essa categoria, de um ponto de vista sincrônico, nas diferentes línguas e provavelmente já na fase mais recente do indo-europeu, passasse a ser em grande parte imotivada semanticamente, mantendo apenas a função de classe de concordância.

Há contudo uma motivação de ordem gramatical que é preciso ainda considerar. Como já se disse, a categoria de gênero organiza o léxico das línguas que a possuem. Essa organização, pelo menos em princípio, parece que estabelecia, no indo-europeu, restrições quanto às relações entre substantivos e verbos, a saber: a) os nomes animados poderiam ocorrer como sujeitos de verbos ativos, como ‘correr’, ‘destruir’, ‘comer’, ‘beber’, ‘morrer’, ‘falar’, ‘matar’, ‘crescer’; b) os nomes inanimados não poderiam exercer a função de sujeitos de verbos ativos, mas apenas de verbos que semanticamente denotassem ações ou estados compatíveis com sujeitos inativos (cf. GRAMKRELIDZE; IVANOV, 1995, p. 239).

Tomemos como exemplo o seguinte enunciado:

I	II	III
sujeito (agente)	verbo (ativo)	objeto (inativo)
o homem	rola	a pedra.

Existe um rol de termos que podem ocupar a posição I (como mulher, animal, tempestade, vento etc.), mas não são todos os itens lexicais que podem fazê-lo: ‘a pedra rola o homem’ seria, por exemplo, um enunciado sem sentido. Em princípio, as palavras que poderiam ocupar a posição I são as que pertenceriam à classe dos animados (ou ativos), não podendo fazê-lo os nomes inanimados (ou inativos). Assim se entende por que o conceito de animado não se restringe a seres que, da nossa perspectiva, têm vida ou se movem, pois seriam perfeitamente coerentes enunciados como:

I	II	III
sujeito (agente)	verbo (ativo)	objeto (inativo)
o céu	fecunda	a terra.
a terra	produz	a árvore.
a árvore	gera	o fruto.

A esse propósito, refletindo um tanto da mentalidade antiga relativamente à natureza, contamos com uma composição das odes anacreônticas (atribuídas, na época helenística, ao poeta grego arcaico Anacreonte), que diz:

A terra negra bebe,
E a árvore bebe a terra.
O mar bebe a brisa,
O sol bebe o mar,
E, ao sol, a lua bebe.
Por que me criticais, camaradas,
Eu também querendo beber?

Ainda que se trate de poesia, em que se trabalha com um registro figurado, o poeta fala de uma experiência de mundo em que terra, sol, lua, mar, árvore são capazes de ocupar a posição de agentes e não só de pacientes.

Um fato de ordem morfológico-sintática que parece corroborar isso é que nos neutros, em geral, a forma do nominativo (o caso do sujeito) é a mesma que do acusativo (o caso do objeto) – noutros termos, trata-se de palavras que originalmente não teriam uma forma de nominativo porque, na condição de inanimados, não poderiam ocupar a posição de sujeitos de verbos ativos. Tudo isso leva Gamkrelidze e Ivanov (1995, p. 239) a concluir:

A divisão de nomes em ativos e inativos, reconstituída para o indo-europeu, encontra paralelos tipológicos frequentes em diversas línguas com uma classificação binária de nomes. A classificação binária motiva a totalidade da estrutura gramatical e os recursos sintático-semânticos do estágio mais remoto que se pode estabelecer para o proto-indo-europeu, a partir da reconstituição comparativa e interna baseada nas línguas indo-europeias. Isso conta para várias das características do indo-europeu, envolvendo as relações gramaticais, sintáticas e semânticas.

Todo esse esforço histórico-comparativo de compreensão da motivação da repartição dos gêneros nas diferentes fases do proto-indo-europeu não elimina o fato de que, com exceção do inglês, nas línguas dessa família se trate de uma categoria em parte imotivada. Repare bem: dizer imotivada parcialmente não implica que o seja de todo. Um falante do português, por exemplo, sabe que o gênero masculino é próprio dos machos, enquanto o feminino designa as fêmeas, compreendendo ambos ainda, em geral, o seguinte:

- (a) são masculinos os nomes de funções exercidas por homens (o cardeal), de rios (o Amazonas), de mares (o Mediterrâneo), dos meses (janeiro vindouro), dos pontos cardeais;
- (b) são femininos os nomes de funções exercidas por mulheres (a freira), de cidades e ilhas (a antiga Ouro Preto, a pacata Paquetá).

Um falante do português sabe também que há razões de ordem fonética ou morfológica para a classificação de gêneros, tais como:

- (a) são masculinas, em geral, as palavras terminadas em *-o* átono (livro, barco, aluno, tesouro), os substantivos concretos terminados em *-ão* (limão, algodão, balcão) etc;
- (b) são femininas, em geral, as palavras terminadas em *-a* átono (*caneta, lancha, aluna, tesoura*), as palavras abstratas terminadas em *-ão* (instrução, razão, aflição), as palavras terminadas em *-agem* (*viagem, garagem, bobagem*) etc (cf. CUNHA, 1982, p. 199-200).

Em todos os casos, os critérios são de ordem bastante genérica e termos que fogem às previsões estabelecidas saltam à vista: ‘cobra’ (gênero feminino) pode ser tanto macha quanto fêmea; o topônimo ‘Rio de Janeiro’ pertence ao gênero masculino; ‘mão’, embora substantivo concreto terminado em *-ão*, é do gênero feminino; ‘cometa’, ‘telefonema’, ‘dia’, ‘fantasma’, ‘mapa’, todos terminados em *-a* átono, pertencem ao gênero masculino; ‘personagem’ admite tanto o masculino (‘o personagem’) quanto o feminino (‘a personagem’). E assim por diante.

Isso leva a que se pense muitas vezes que o gênero é imotivado, o que não seria de todo correto. O mais adequado, tendo em vista a variedade de línguas do mundo, parecer ser considerar que há gradações, admitindo-se tanto critérios semânticos, quanto morfológicos, na seguinte escala, de acordo com Corbett (1991): (a) sistemas estritamente semânticos; (b) sistemas predominantemente semânticos; (c) sistemas morfológicos; (d) sistemas fonológicos – prevendo-se ainda que possa haver a combinação de critérios semânticos, morfológicos e fonológicos, o que se aplicaria, em geral, às línguas indo-europeias, com a eventual predominância de um ou outro, como no russo, em que prevalece a morfologia (cf. CORBETT, 1991, p. 1-61).

O inglês é a única língua indo-europeia que se enquadraria no primeiro item acima citado, ou seja, o dos “sistemas estritamente semânticos”, com três gêneros: masculino, feminino e neutro. Como o artigo, os demonstrativos e os adjetivos não apresentam variação de gênero, não concordando, portanto, com o nome que determinam, há linguistas que consideram que essa categoria gramatical simplesmente não existe no inglês moderno. Todavia, o gênero se encontra expresso nos pronomes de terceira pessoa do singular (*he/she/it*) e nos possessivos a eles correspondentes (*his/her/its*), exigindo a concordância com o referente, de acordo com a seguinte distribuição: a) *he* (masculino), humanos machos; b) *she* (feminino), humanos fêmeas; c) *it* (neutro), todo o restante. As exceções são raras, como *ship*, ‘navio’, que frequentemente admite o pronome *she*. Entretanto, essa divisão estritamente semântica permanece aberta, na prática, a uma maior variedade: animais domésticos, sobretudo quando têm nomes próprios, admitem o uso de *he* ou *she*, dependendo de seu sexo, o mesmo acontecendo nas histórias infantis, ou seja, fatores pragmáticos, de ordem emotiva e cultural, interferem na escolha do falante, nas diversas situações comunicativas.

2.3 A categoria de caso

Podemos definir *caso* em dois sentidos: (a) como as funções que uma palavra assume numa estrutura sintática (sujeito, objeto etc); (b) como a forma assumida por determinada palavra, que a vincula ao exercício de determinadas funções sintáticas (nominativo, acusativo etc).

Neste último sentido, que é o que nos interessa do ponto de vista morfológico, registram-se duas maneiras de marcação de caso:

- (a) Sob a forma de *declinação*, ou seja, pelo acréscimo de afixos à palavra;
- (b) Pela existência de séries supletivas.

Assim, em português os pronomes pessoais apresentam diferenciação de caso pelo uso de formas supletivas, de acordo com a seguinte distribuição:

QUADRO 12

Formas supletivas dos pronomes pessoais em português nos casos reto e oblíquo				
Pessoas		Caso reto (função de sujeito)	Caso oblíquo (demais funções)	
			Formas átonas	Formas tônicas
singular	1ª. pessoa	eu	me	mim/eu
	2ª. pessoa	tu/você	te/o/a/lhe	ti/você/tu
	3ª. pessoa	ele/ela	o/a/lhe	ele/ela
plural	1ª. pessoa	nós/a gente	nos	nós/a gente
	2ª. pessoa	vós/vocês	vos/lhes	vós/vocês
	3ª. pessoa	eles/elas	os/as/lhes	eles/elas

A primeira pessoa apresenta a distribuição mais conservadora, mantendo as duas raízes existentes também em outras línguas indo-européias, a saber:

- (a) Para o caso reto (ou nominativo, função de sujeito), *eu* < nominativo latino *ego* < i.e. **egho(m)*;
- (b) Para o caso oblíquo, *me* < acusativo latino *me* < i.e. **me*, bem como *mim* < português arcaico *mi* < dativo latino *mihi* < i.e. **me*.

A forma “eu” é exclusiva do sujeito (‘eu corro/eu te amo’), sendo que seu uso como objeto direto parece restrito a registros peculiares (há alguns anos fez sucesso uma canção que dizia “leva eu, minha saudade”, correspondente a “me leva, minha saudade”). Uma oração como ‘ela gosta de eu’ seria sentida como agramatical ou, pelo menos, não usual por falantes do português, pois se esperaria ‘ela gosta de mim’. A alternância entre a função de sujeito (nominativo) e os demais casos se expressa bem na oposição entre ‘eu te amo’/‘ela me ama’/‘ela gosta de mim’.

Observando-se o quadro acima, percebe-se que há, no português contemporâneo, possibilidades variadas, algumas compreendendo a neutralização das séries supletivas, como, por exemplo, com relação à segunda pessoa. Imagine que estamos dizendo a Maria que Pedro a ama. As possibilidades, nas diversas situações possíveis no diassistema do português brasileiro, seriam: 'ele te ama'/'ele a ama'/'ele ama você'/'ele lhe ama'. Mais estranha e, em certa medida, percebida como agramatical, seria uma construção como 'ele ama tu', embora, também no campo da música popular, tenha feito sucesso uma canção que dizia "passei a noite procurando tu, procurando tu, procurando tu", em que o uso de 'tu' (caso reto/nominativo) como objeto direto (caso oblíquo/acusativo) produzia efeito justamente pelo que teria de inusual e pela rima que a repetição do sintagma propiciava, reforçando-o.

É preciso considerar que esse tipo de neutralização se observa mesmo em línguas que contam com um sistema de flexão de casos (declinação), o nominativo podendo, ainda que em exemplos que são um tanto raros, ocupar posições em que se esperariam outros casos. Isso indica que, em geral, se trata de um caso não marcado. Se, em português, 'eu passei a noite procurando tu' é uma construção possível, o mesmo não se diria de *'me passei a noite procurando tu'. Do mesmo modo, se 'ele ama ela' é uma forma corrente, alternando com 'ele a ama', *'o a ama' é inadmissível e 'o ama ela' equivale a 'ela ama ele' ou 'ama-o ela', em que 'ela' chama para si a função de sujeito, mesmo estando depois do verbo.

Os últimos exemplos mostram que, ocorrendo neutralização das formas supletivas, a posição dos termos da oração torna-se responsável pela marcação de caso, situando-se o sujeito, em português, antes do verbo e o objeto depois, ou seja, há uma ordem estabelecida Sujeito-Verbo-Objeto (SVO): assim, 'ele ama ela' (= 'Pedro ama Maria') é diferente de 'ela ama ele' (= 'Maria ama Pedro'), sendo marcado como sujeito o termo que se põe antes do verbo e como objeto o que aparece depois. Não ocorrendo a neutralização, a ordenação dos termos torna-se menos rígida. Assim, em 'ela lhe disse palavras de amor'/'disse-lhe ela palavras de amor'/'lhe disse ela palavras de amor', o pronome 'ela' (caso reto/nominativo) mantém sempre a função de sujeito, enquanto 'lhe' (caso oblíquo/dativo) conserva também sempre a função de objeto indireto, com base no contraste provido pelas duas formas supletivas.

Tomemos mais um exemplo. Em francês, diferentemente do que acontece em português, o pronome relativo apresenta duas formas, a saber: *qui* (nominativo < nominativo latino *qui*) e *que* (acusativo < acusativo latino *quem*). Portanto, enunciados como 'o homem que me viu é meu amigo' (em que o pronome 'que' exerce a função de sujeito, já que foi ele que me viu) e 'o homem que eu vi é meu amigo' (em que o pronome 'que' tem a função de objeto direto, já que fui eu que vi o homem) se dirão assim:

QUADRO 13

Formas de nominativo e acusativo do pronome relativo em francês

Função do pronome relativo	Pronome	Exemplos
Sujeito	<i>qui</i>	<i>l'homme qui m'a vu est mon ami</i> / o homem que me viu é meu amigo
Objeto direto	<i>que</i>	<i>l'homme que j'ai vu est mon ami</i> / o homem que eu vi é meu amigo

2.3.1 A categoria de caso no indo-europeu

A categoria gramatical de caso parece ter sido morfológicamente marcada através de sufixos no indo-europeu, pelo menos no estágio até onde é possível reconstituir essa protolíngua. Como sempre, isso se deduz pela situação das diferentes línguas que dela provêm, admitindo essas, numa distribuição bastante variada, desde a ausência de marcações de casos, quanto a flexão do substantivo e de seus determinantes em até oito casos, a saber:

- (a) Nominativo: é o caso próprio do sujeito, como no seguinte exemplo do latim: *Petrus amat Mariam* (Pedro ama Maria), em que a terminação *-us* indica que *Petrus* exerce a função de sujeito do verbo *amat* (ama), independentemente da ordem dos termos, pois, sem alteração de sentido, se poderia dizer *Petrus Mariam amat/Mariam Petrus amat/Mariam amat Petrus* etc.;
- (b) Acusativo: o caso do objeto direto, ou seja, do complemento de um verbo transitivo, papel exercido no exemplo acima por *Mariam*, que recebe a terminação *-am* por exercer a função de objeto direto – caso fosse Maria que amasse Pedro, então os papéis se inverteriam e *Petrus* assumiria a forma do acusativo: *Maria Petrum amat* (Maria ama Pedro);
- (c) Dativo: marca o objeto indireto, ou seja, o beneficiário da ação expressa pelo verbo, como, ainda em latim, *Maria librum Petro dat* (Maria dá um livro para Pedro), em que a terminação *-a* indica que *Maria* é o sujeito, a terminação *-um*, que *librum* é o objeto direto, e a terminação *-o*, que *Petro* constitui o objeto indireto, a ordem dos termos da oração podendo variar sem alteração do sentido;
- (d) Genitivo: indica posse ou o determinante nominal de outro nome, como nos exemplos, também do latim, *domus Petri* (casa de Pedro)/*domus Mariae* (casa de Maria), ou, no inglês, *Peter's house* (casa de Pedro) – o que equivale à anteposição ao nome das preposições 'de' (em português, espanhol, francês etc), *di* (italiano), *of* (inglês), *von* (alemão) etc;
- (e) Ablativo: indica procedência, como, de novo em latim, *Roma venio* (venho de Roma) – essa relação é expressa em português, espanhol e francês também pela preposição 'de' (venho do Rio de Janeiro/*vengo de Madrid/je viens de Paris*), mas, em italiano, por

uma preposição diferente, ou seja, *da* (*vengo da Roma*), do mesmo modo que, em inglês, por *from* (*I come from New York*) e, em alemão, por *aus* (*Ich komme aus Berlin*) etc;

- (f) Locativo: indica o lugar onde, sem ideia de movimento, como, em russo, *dieti v shkolie* (as crianças [estão] na escola) – no português, essa função se expressa com a preposição *em*;
- (g) Instrumental-associativo: cobre a esfera expressa, em português, pela preposição *com* (eu escrevo com a caneta/eu passeio com meu amigo), como no exemplo do russo, *ia nishu karandashom* (eu escrevo com o lápis);
- (h) Vocativo: marca o chamamento, a invocação, como, em latim: *Petre, amas Mariam?* (ó Pedro, amas Maria?).

No quadro abaixo, você encontra indicados os casos em que os nomes e seus determinantes são regularmente flexionados em algumas línguas indo-europeias de diferentes grupos, embora não todas apresentem flexão para todos esses casos, como se mostra no quadro abaixo:

QUADRO 14

Casos morfológicamente marcados, através de flexão, em algumas línguas indo-europeias

	Nom.	Acus.	Gen.	Dativo	Ablat.	Locat.	Instr.	Vocat.
Hitita	x	x	x	x	x		x	
Sânscrito	x	x	x	x	x	x	x	x
Persa antigo	x	x	x	x	x	x	x	x
Armênio	x	x	x	x	x	x	x	
Lituano	x	x	x	x		x	x	x
Russo	x	x	x	x		x	x	
Albanês	x	x	x	x	x			
Grego antigo	x	x	x	x				x
Grego moderno	x	x	x					
Alemão	x	x	x	x				
Inglês			x					
Irlandês	x	x	x	x				x
Latim	x	x	x	x	x			x
Romeno	x	x	x	x				x

Esse levantamento é bastante representativo na medida em que as línguas tomadas como exemplos pertencem a diferentes ramos do indo-europeu: anatólio (com o hitita), indo-iraniano (com o sânscrito e o persa antigo), báltico (com o lituano), eslavo (com o russo), germânico (com o alemão e o inglês), celta (com o irlandês), itálico (com o latim e o romeno), armênio, albanês e grego. Entretanto, quanto mais se avança para a esquerda menos regular

se torna a marcação morfológica de caso, com no exemplo extremo do inglês que conserva morfológicamente marcado apenas o genitivo.

No quadro abaixo, que toma como referência o pronome demonstrativo masculino, você poderá observar como a marcação se processa:

QUADRO 15

Declinação do pronome demonstrativo 'esse' (em latim) e 'este' (nas demais línguas)

Casos	Sânscrito	Russo	Latim	Grego antigo	Alemão
Nominativo (<i>este/esse</i>)	<i>eshaḥ</i>	<i>etot</i>	<i>iste</i>	<i>hoûtos</i>	<i>dieser</i>
Acusativo (<i>este</i>)	<i>etam</i>	<i>etogo</i>	<i>istum</i>	<i>toûton</i>	<i>diesen</i>
Genitivo (<i>deste</i>)	<i>etasya</i>	<i>etogo</i>	<i>istius</i>	<i>toútou</i>	<i>dieses</i>
Dativo (<i>a este</i>)	<i>etasmai</i>	<i>etomu</i>	<i>istī</i>	<i>toútōi</i>	<i>diesem</i>
Ablativo (<i>deste</i>)	<i>etasmāt</i>		<i>istō</i>		
Locativo (<i>neste</i>)	<i>etamin</i>	<i>etom</i>			
Instrumental (<i>com este</i>)	<i>etena</i>	<i>etim</i>			

Você pode se perguntar como procedem as línguas que não apresentam flexão para todos os casos, a resposta sendo muito simples: apelam elas para o uso de preposições. No sânscrito não há necessidade de preposições, pois os sete casos se encontram sistematicamente marcados. Nas demais línguas a situação relativa aos casos que não têm flexão própria apresenta-se assim: (a) utilizam-se preposições; (b) cada preposição pede um dos casos flexionados, o que se chama regência da preposição:

QUADRO 16

Preposições que suprem a flexão de casos e sua regência em grego, russo, alemão e latim

Regência das preposições	Preposições que expressam o sentido dos casos não flexionados								
	Ablativo (lugar de onde)			Locativo (lugar onde)			Instrumental (com quê)		
	grego	russo	alemão	grego	alemão	latim	grego	alemão	latim
Genitivo	<i>ek/apó</i>	<i>iz/ot</i>							
Dativo			<i>aus</i>	<i>en</i>	<i>in</i>		<i>sún</i>	<i>mit</i>	
Ablativo						<i>in</i>			<i>cum</i>

Há duas formas de considerar essa situação: (a) na primeira, própria da reconstituição tradicional do indo-europeu, conforme a formulação de Brugmann, considera-se que a protolíngua apresentaria os oito casos conservados em sânscrito e o processo de mudança nas línguas dela derivadas seria constituído por uma simplificação da flexão nominal; (b) a segunda, assumida por autores como Rodríguez, Villar, Gamkrelidze e Ivanov, ensaia matizar as diferentes fases do indo-europeu desde um possível estágio não-flexionado, para

tentar entender como se processa a criação da flexão nominal de caso. É esta última hipótese que seguiremos.

2.3.2 A criação das oposições de caso

Partindo-se da hipótese bastante razoável de que o indo-europeu tenha conhecido uma fase não flexionada, as primeiras oposições morfológicamente marcadas de caso, ou seja, a primeira flexão criada para os nomes, com o acréscimo de desinências ao radical, parecem ter sido as relativas ao genitivo e ao acusativo. Isso se constata, antes de tudo, pela existência de terminações para os dois casos que se registram nos vários grupos da família indo-europeia, o que testemunha em favor de sua antiguidade. Essas terminações seriam **(o)s* para o genitivo e **(o)m* para o acusativo. Importante é observar que, com relação ao primeiro, se trata de estabelecer a relação entre o determinante e o determinado no sintagma nominal ('casa de Pedro' ou 'casa de pedra'), enquanto com relação ao segundo a relação entre determinante e determinado diz respeito ao sintagma verbal ('vi Pedro' ou 'vi a casa'). Noutros termos: é razoável admitir que a flexão se tenha iniciado para marcar as relações sintáticas nesses dois pontos mais básicos dos enunciados.

Antes da criação do sistema de flexão nominal, é natural supor que o indo-europeu contasse com outros recursos para marcar as referidas funções, como a ordem dos termos, a entonação, a variação do acento etc. Em especial a ordem dos termos se mostra bastante eficaz para exprimir relações sintáticas, como você sabe que acontece em português com relação ao verbo e o objeto (VO). No que diz respeito ao sintagma nominal, também a ordem dos termos pode ser significativa por si, como nos exemplos, também do português, 'palavra-chave', 'cidade-estado', 'homem-bomba', 'operação limpeza' (ordem Determinado-Determinante), em que apenas o fato de um substantivo vir justaposto e após outro já cria a relação de determinação (observe-se que nem mesmo é necessário proceder a concordância de número: 'palavras-chave', 'cidades-estado' etc).

No caso do indo-europeu, ainda que seja impossível saber com certeza os recursos que permitiriam expressar as relações sintáticas antes da criação da flexão, a ordem dos termos parece ter sido significativa, com o determinante precedendo o determinado, no caso dos sintagmas nominais, como sugerem alguns nomes compostos de diversas línguas:

QUADRO 17

Nomes compostos e ordenação dos elementos envolvidos na composição

Língua	Composto	Significado	Elementos de composição	
			Determinante	Determinado
Sânscrito	<i>mahādevā-</i>	'grande deus', 'Shiva'	<i>mahā-</i> , 'grande'	<i>devā-</i> , 'deus'
	<i>svāsthāna-</i>	'terra natal'	<i>svā-</i> , 'seu, sua'	<i>sthāna-</i> , 'terra'
Grego	<i>akrópolis</i>	'cidadela, cidade alta'	<i>ákro-</i> , 'alto'	<i>pólis</i> , 'cidade'
	<i>rhododáktulos</i>	'de dedos róseos'	<i>rhodo-</i> , 'rosa'	<i>dáktulos</i> , 'dedo'
Gótico	<i>armahairts</i>	'misericordioso'	<i>arma-</i> , 'piedade'	<i>hairts</i> , 'coração'
Latim	<i>misericiórdia</i>	'misericórdia'	<i>miser-</i> , 'desventurado'	<i>cord-</i> , 'coração'
Russo	<i>polovod'e</i>	'enchente'	<i>pol-</i> , 'cheio'	<i>vod'-</i> , 'água'

2.3.2.1 Emergência do "relacionador" do sintagma nominal (genitivo)

Nesse contexto em que os itens lexicais se organizam nos sintagmas sem marcas morfológicas de classe ou de caso, sendo provável que não existisse diferença entre substantivos e adjetivos (o adjetivo sendo o determinante nominal por excelência), a terminação **-os/es/s* começa a ser usada para relacionar dois nomes, constituindo pares de oposição entre um caso não marcado (o do determinado) e um caso marcado por sufixo (o do determinante), como nos seguintes exemplos de genitivos: grego *pod-ós*, latim *ped-is*, sânscrito *pad-ás* (*-as* < **-os*), 'do pé'; hitita *nepiš-aš* (*-aš* < **-os*), 'do céu'. Parece que essa terminação aplicava-se tanto ao singular quanto ao plural, antes que fosse criada uma flexão própria de número, tanto que, em hitita, encontramos, como genitivo singular e plural de *uddar*, 'palavra', o termo flexionado *uddanaš*.

Ora, é curioso que haja ainda uma segunda terminação de genitivo atestada nas línguas indo-europeias: **-ōm*. Também esta parece servir tanto para o singular quanto para o plural, como encontramos, ainda em hitita, *antuhšan* (*-an* < **-ōm*), 'de uma pessoa', ao lado do latim *pedum* (< **pedōm*), 'dos pés', e do antigo eslavo eclesiástico *imenŭ* (*-ŭ* < **-ōm*), 'dos nomes'. Parece que, inicialmente, a distribuição entre as duas terminações deveria ter relação com o gênero de cada palavra, os nomes animados recebendo a terminação **-os* e os inanimados a terminação **-ōm*.

Posteriormente, com o acréscimo a esta última terminação (**-ōm*) da desinência de plural **-s* (**-oms* > **ōm*), a mesma se teria especializado para expressar a ideia apenas de genitivo plural, como encontramos no grego *pedōn*, 'dos solos', sânscrito *padām*, 'dos pés, dos passos', lituano *vilku* (*-u* < **-ōm*), 'dos lobos' (cf. GRAMKRELIDZE; IVANOV, 1995, p. xx).

Declinar uma palavra é, no fundo, mudar-lhe a classe. Assim, os substantivos na forma do genitivo (expressando uma relação que, em português, estabelecemos com a preposição

'de') correspondem a adjetivos, classe de palavras que se define enquanto determinantes de nomes. Quando dizemos 'coração de pedra', 'de pedra' constitui uma locução adjetiva, tanto que poderia ser substituído por um verdadeiro adjetivo, 'coração pétreo'. É razoável supor, como o faz Villar Liebana (1974, p. 251-262), que os nomes flexionados no genitivo estejam na origem da classe dos adjetivos do indo-europeu: quando o grego opõe, por exemplo, o nominativo *kuon* (cão) ao genitivo *kunós*, esta última palavra assume uma função adjetiva ('mordida de cão' ou 'mordida canina').

2.3.2.2 A emergência da marca de determinante do sintagma verbal (acusativo)

Como você já sabe, o acusativo é o caso próprio do objeto. Portanto, quando um nome recebe o sufixo próprio deste caso, isso indica tratar-se de um determinante do verbo, como quando digo 'eu como pão', em que 'pão' determina o verbo 'comer' (não se trata de comer 'carne', 'frutas' ou seja lá o que for, mas 'pão').

Parece que, paralelamente à criação do morfema de genitivo, o estabelecimento de uma marca morfológica para o acusativo se deu numa fase muito arcaica do indo-europeu, anterior a sua separação em diferentes dialetos e línguas. O indício mais importante para chegar-se a essa conclusão é que o acusativo apresenta, nos vários ramos do indo-europeu, uma terminação bastante regular, com uma consoante nasal (-*m/-n*), como se pode constatar no quadro seguinte:

QUADRO 18

Exemplos da oposição nominativo/acusativo

Casos	Sânscrito	Grego	Latim	Gótico	Hitita
Nominativo	<i>vṛkaha</i> (lobo)	<i>lúkos</i> (lobo)	<i>lupus</i> (lobo)	<i>hana</i> (galo)	<i>ginuš</i> (joelho)
Acusativo	<i>vṛkama</i>	<i>lúkon</i>	<i>lupum</i>	<i>hannan</i>	<i>ginun</i>

Se você não esqueceu que o gênero neutro corresponderia a palavras do tipo inativo, ou seja, que não poderiam exercer a função de sujeito de verbos ativos, é importante agora observar que esses nomes, no sânscrito, no grego, no latim e no hitita, têm como característica, na declinação temática, o fato de que o nominativo termine com uma *-m* ou *-n*, bastando observar de novo a distribuição de gênero nos adjetivos:

QUADRO 19
Gêneros nos adjetivos

	Masculino	Feminino	Neutro
Sânscrito	<i>kântah</i> (amado)	<i>kântâ</i>	<i>kântam</i>
Grego	<i>ponerós</i> (infeliz)	<i>ponerá</i>	<i>ponerón</i>
Latim	<i>bonus</i> (bom)	<i>bona</i>	<i>bonum</i>

Assim, podemos dizer que, na origem, acrescentar a um termo ativo o sufixo de acusativo equivaleria a mudá-la de classe, para indicar que, naquela situação, ela passara a exercer a função sintática própria dos nomes inativos, ou seja, a de objeto.

2.3.2.3 Os outros casos

No que diz respeito aos outros casos, é difícil descobrir padrões comuns que se apliquem ao conjunto do indo-europeu, parecendo antes que os elementos morfológicos são mais recentes, tendo sido criados já numa fase de dialeção, cada grupo gramaticalizando certos morfemas para expressar outras relações sintáticas.

É razoável supor ainda que alguns ramos do indo-europeu levaram mais longe a tendência a criar marcas nominais de caso, como o indo-iraniano e o eslavo, outros fazendo-o menos, como o germânico e o grego. No caso desta última língua, é provável que alguns sufixos locativos estivessem em vias de gramaticalizar-se como autênticas desinências, como os sufixos *-then* para expressar o ablativo (*lugar de onde*) e *-de* para o alativo (*lugar para onde*) – como em *oíkothen* (*de casa*) e *oíkade* (*para a casa*) –, o que todavia não aconteceu, mantendo-se eles apenas como formadores de alguns advérbios.

Guia de Leitura Texto 4: A Reconstrução do Indo-Europeu

- (1) Quais são os quatro itens que a reconstrução de uma protolíngua pretende estabelecer?
- (2) O que significa dizer que a protolíngua deve ser *tipologicamente viável*? Ilustre sua explicação com um exemplo.
- (3) Qual o primeiro passo para a reconstrução de uma língua a partir do método comparativo?
- (4) Qual a evolução do nexos consonantal *-ct* nas línguas advindas do latim?
- (5) Em que consistia o chamado *Apendix Probi*? Em que esta obra pode ser útil para a apuração da evolução histórica de palavras como *orelha*?
- (6) Conceitue e exemplifique o fenômeno fonético conhecido por *rotação germânica*.
- (7) Qual a organização hierárquica que regula as três noções de número nas línguas indo-europeias (singular, plural e dual)?
- (8) Cite duas línguas indo-europeias em que o dual esteja presente.
- (9) Em que fase evolutiva supõe-se ter ocorrido a gramaticalização do dual no indo-europeu?
- (10) Em que situações usa-se o dual nas línguas indo-europeias que dispõem desse recurso gramatical? A que referentes, no mundo físico, o dual tipicamente está associado?
- (11) Quando a categoria de gênero pode ser percebida como gramatical?
- (12) Quanto à categoria de gênero, as línguas indo-europeias ramificam-se em cinco grupos, cada qual marcado por um número específico de distinções de gênero. Quais são estes cinco grupos?
- (13) Quais as ordens de motivação semântica envolvidas nas distinções de gênero das línguas indo-europeias?
- (14) O que a existência do modelo tripartido de gêneros (masculino/feminino e neutro) no avéstico, no sânscrito, no grego e no latim sugere com relação ao indo-europeu?
- (15) Por que muitos nomes relativos aos corpos celestes e a elementos da natureza, como o céu ou a terra, receberam marcas de feminino e masculino e não uma marca de neutro no indo-europeu?
- (16) Considerando-se o modelo tripartido, quais nomes podem exercer a função de sujeitos de verbos ativos (nomes masculinos, femininos ou neutros)?
- (17) Quais as duas definições possíveis para a categoria de *caso*?
- (18) Como se dá a marcação de caso por meio de *séries supletivas*? Exemplifique.
- (19) Dê a definição dos seguintes casos: nominativo, acusativo e dativo.
- (20) Dê a definição dos seguintes casos: genitivo, ablativo e locativo.
- (21) Dê a definição dos seguintes casos: instrumental-associativo e vocativo.
- (22) Qual o procedimento utilizado por línguas que não possuem flexão para todos os casos para expressar o conteúdo semântico dos casos para os quais não dispõe de marcação morfológica?

TEXTO 5

O QUE É UMA LÍNGUA

Tommaso Raso e Jacyntho Lins Brandão

Não é possível saber com exatidão quantas línguas existem no mundo, as estimativas oscilando entre cinco e dez mil. Estabelecer essa quantidade depende, dentre outras coisas, dos critérios com que se distinguem línguas e variedades de uma mesma língua. Por exemplo, o português de Portugal e o português do Brasil são considerados duas variedades da mesma língua, com base no fato de que há compreensão recíproca. Utilizando-se o mesmo critério, contudo, os dialetos de Nápoles e de Milão deveriam ser tidos como línguas diferentes, porque não há compreensão recíproca.

Caso se adote o critério de inteligibilidade mútua, a estimativa é de que são faladas atualmente mais de seis mil línguas – conforme a última edição do *Ethnologue*, com data de 2009, exatamente 6.909 (Lewis, *Ethnologue*; dados disponíveis também no site www.ethnologue.com).

1 Línguas nacionais, regionais e minoritárias

Pode-se perguntar por que é tão complicado estabelecer o número de línguas faladas no mundo. Uma das dificuldades está em que há sempre fatores de ordem histórica, identitária, cultural e política que interferem no reconhecimento social de uma língua. Sobretudo a partir do século XVIII, com a formação dos estados nacionais, predominou a ideia de que uma nação se define por um território e uma única língua, a chamada “língua oficial”, ensinada nas escolas e admitida em documentos públicos, assunto geralmente tratado como tema constitucional.

Há nessa concepção dois tipos de problema: (a) por um lado, os que surgem em vista do fato de que no interior de uma mesma fronteira geográfica muitas vezes se encontra mais de uma língua (o que se constata exemplarmente no caso da Espanha, onde se falam o galego, o espanhol, o catalão e o basco, para citar apenas as que contam, além das variantes orais, também com uma rica literatura; ou no caso do Peru, com o espanhol, o quêchua e o aimara); (b) por outro lado, levanta questões ainda o fato de que a abrangência geográfica de uma língua pode ultrapassar as fronteiras nacionais (o basco, por exemplo, é falado num território contínuo que se estende tanto pela Espanha, quanto pela França, bem como o espanhol, na

América, abrange um grande número de nações, o mesmo sendo verdadeiro com relação ao quêchua, falado tanto no Peru, quanto na Bolívia, na Colômbia, no Equador, no Chile e na Argentina). A concepção de que cada nação deve ter apenas uma língua, responsável pelo desprestígio e até pela extinção de muitas línguas minoritárias, só pouco a pouco e com muitos empecilhos tende a se modificar, a partir da perspectiva de que a diversidade linguística representa riqueza e não um problema. Não se trata, entretanto, de algo simples.

Comprova-o caso do bretão, falado no norte da França. Até poucas décadas atrás, nas repartições públicas e nas escolas havia cartazes com os dizeres: *il est interdit de cracher par terre et de parler breton* (“é proibido cuspir no chão e falar bretão”). Isso implicou que essa língua se restringisse praticamente a um uso doméstico, sendo em geral deixada de lado pelos homens, em vista de sua integração nos espaços públicos, e mantida majoritariamente pelas mulheres, então mais restritas à esfera privada. Atualmente, a Comunidade Europeia incentiva a diversidade linguística, tendo mesmo sido aprovado, em 1992, pelo Conselho Europeu, o “Estatuto Europeu das Línguas Regionais ou Minoritárias” (ETS 148), o qual prevê, da parte dos diversos países, o reconhecimento das línguas minoritárias faladas em seu território, bem como uma série de medidas visando a sua manutenção, tais como sua presença no sistema de ensino para famílias e estudantes que assim o desejarem, a publicação nelas de textos oficiais, o apoio a sua utilização nos meios de comunicação etc (cf. Conselho da Europa, *European Charter for Regional or Minority Languages*). A maioria dos membros da Comunidade Europeia já ratificou o estatuto, mas, na França, a Comissão pertinente da Assembleia Nacional considerou, em junho de 1999, que ele contraria o artigo 2 da Constituição, o qual prevê simplesmente que “a língua da República é o francês”. Desse modo, conforme o deputado bretão François de Ruy, “cada vez que um dos deputados (incluindo eu próprio) tem o azar de pronunciar apenas uma palavra numa das nossas línguas regionais, o Presidente da sessão se apressa em lembrar que isso é proibido e que o dito não será reproduzido na ata” (cf. Ruy, *À l’Assemblée, il est interdit de parler breton...* [Na Assembleia Nacional é proibido falar bretão...]). Saliente-se que a França não é um caso isolado: também Bélgica, Grécia, Irlanda, Portugal, Itália e Rússia, dentre outros, ainda não ratificaram o citado estatuto.

Um outro exemplo, se bem que em sentido contrário, demonstra o quanto língua e política podem manter relações bastante intrincadas: o do “moldavo”. Nenhum linguista duvida de que se trata não de uma língua, mas de um dialeto do romeno. Embora este último tenha recebido, antes do século XVIII, denominações variadas, como valáquio e moldavo, já em 1574 o viajante francês Pierre Lescalopier observava que os habitantes da Moldávia, da Valáquia e da Transilvânia constituíam uma unidade etnolinguística que ele denominou “*românesc*” (romena). Em 1716, também Dimitrie Cantemir, um dos grandes iniciadores da

cultura romena, em sua *Hronicul vechimei romano-moldo-vlahilor* (Crônica da antiguidade romeno-moldavo-valáquia), defendia uma configuração em que a unidade romena tinha como base os moldávios e os valáquios. Mas, ao mesmo tempo, falava tanto de uma “língua moldava” (*limba moldovenească*) quanto de uma “língua romena” (*limba româniască*), ainda que ressaltando que a primeira se parecia com a segunda mais que qualquer outra. A denominação “língua moldava” permaneceu e passou a ser oficialmente utilizada quando, em 1812, a Moldávia foi anexada ao Império russo e rebatizada como “província da Bessarábia”; nesse mesmo ano, o moldavo foi admitido, pelo tsar Alexandre I, nos procedimentos jurídicos – e em 1818 foi finalmente declarado língua oficial daquela província, ao lado do russo. Isso não impediu que, no próprio contexto do Império russo, o termo “romeno” continuasse a ser utilizado nos meios especializados, como, por exemplo, no título da obra *Cursulu primitiviu de limba rumâna* (*Curso inicial de língua romena*), de I. Doncev, aparecida em 1865 e destinada a uso nas escolas da Moldávia (cf. HANGANU, 2009, p. 89-90).

E a história prossegue, no século XX e no nosso, sempre mesclada aos acontecimentos políticos da região, num vai-e-vem que parece interminável. Em 1918, a República Democrática da Moldávia, proclamada no ano anterior na Bessarábia, juntou-se à Romênia, e a língua falada por seus habitantes foi denominada “romeno” – e não “moldavo”, como faziam os etnógrafos russos. Quando a Rússia anexou novamente a Bessarábia, em 1940, na esteira do acordo entre Hitler e Stalin, o termo “moldavo” voltou a ser oficial na República Socialista Soviética da Moldávia – a posição soviética sendo que se tratava de uma língua diferente do romeno, muito menos latinizada e muito mais influenciada por elementos eslavos, a que se ajunta o fato de que, em 1938, ela passou a ser escrita com o alfabeto cirílico russo (лимба молдовеняскэ). Em 1989, o moldavo foi declarado língua oficial da República da Moldávia, mas o uso do alfabeto latino foi restabelecido. Em 12 de maio de 1990, o moldavo foi oficialmente reconhecido como “romeno”, o que serviu de pretexto para a guerra civil que resultou na criação da República Moldava do Dniestr (ou Transnístria), a qual possui três línguas oficiais: o russo, o ucraniano e o moldavo (escrito no alfabeto cirílico). Por ocasião da independência da República da Moldávia, em agosto de 1991, sua constituição estabeleceu que “a língua oficial da República da Moldávia é a língua romena, e utiliza o alfabeto latino”, o que provocou a reação dos não-romenófilos e, assim, em 1993, voltou-se à denominação de “moldavo”. Em 1996, uma proposta do presidente do parlamento, Mircea Segur, de retomar a denominação de “romeno” foi rejeitada, tendo o governo da Moldávia patrocinado, em 2003, a publicação de um dicionário moldavo-romeno, visando a provar que os dois países falam línguas diferentes. Ressalte-se que a própria Academia de Ciências da Moldávia entende que não se trata de duas línguas, mas de uma única língua romena, com variantes regionais.

Para que se tenha ideia da diversidade linguística e do quanto fronteiras nacionais não

implicam em fronteiras linguísticas, examine-se, no quadro 1, a relação dos países europeus que já ratificaram o citado Estatuto das Línguas Regionais ou Minoritárias:

QUADRO 1

Países que já ratificaram o ETS 148 e as línguas reconhecidas

País	Ano da ratificação	Língua ou línguas oficiais	Línguas reconhecidas como co-oficiais (LCO) e línguas regionais ou minoritárias reconhecidas (LRM)
Alemanha	1998	Alemão	LRM: Baixo-alemão, Dinamarquês, Frísio, Români, Sorábio
Armênia	2002	Armênio	LRM: Aramaico, Grego, Curdo, Russo, Iezidi
Áustria	2001	Alemão	LRM: Croata, Húngaro, Români, Eslovaco, Esloveno
Chipre	2002	Grego, Turco	LRM: Árabe Cipriota Maronita, Armênio
Croácia	1997	Croata	LRM: Húngaro, Italiano, Ruteno, Sérvio, Eslovaco, Esloveno, Tcheco, Ucrâniano
Dinamarca	2000	Dinamarquês	LRM: Alemão
Eslováquia	2001	Eslovaco	LRM: Alemão, Búlgaro, Croata, Húngaro, Polonês, Români, Ruteno, Tcheco, Ucrâniano
Eslovênia	2000	Esloveno	LRM: Alemão, Croata, Húngaro, Italiano, Români
Espanha	2001	Espanhol	LCO: Aranês (Catalunha), Basco (País Basco e Navarra), Catalão (Catalunha e Ilhas Baleares), Galego (Galícia), Valenciano (Valência); LRM: Árabe, Aragonês, Asturiano, Bérbere, Português
Finlândia	1994	Finlandês	LCO: Sueco (Ostrobotnie e Åland); LRM: Iídiche, Lapão, Români, Russo, Sueco, Tártaro
Hungria	1995	Húngaro	LRM: Alemão, Armênio, Beas, Búlgaro, Croata, Eslovaco, Esloveno, Grego, Polonês, Romeno, Români, Ruteno, Sérvio, Ucrâniano
Lichtenstein	1997	Alemão	
Luxemburgo	2005	Alemão, Francês, Luxemburguês	
Montenegro	2006	Montenegrino, Sérvio	LRM: Albanês, Români
Noruega	1993	Norueguês	LCO: Kven, Lapão, Români
Países Baixos	1996	Holandês	LRM: Baixo-Saxão, Frísio, Luxemburguês, Români, Iídiche
Polônia	2009	Polonês	LRM: Alemão, Armênio, Bielorusso, Tcheco, Hebraico, Iídiche, Caraíta, Cabucho, Lituano, Lemcoviano, Români, Russo, Eslovaco, Tártaro, Ucrâniano
Reino Unido	2001	Inglês	LRM: Córnico, Escocês, Gaélico Escocês, Galês, Irlandês, Manês
República Tcheca	2006	Tcheco	LRM: Alemão, Eslovaco, Polonês, Români
Romênia	2007	Romeno	LRM: Albanês, Alemão, Armênio, Búlgaro, Croata, Eslovaco, Grego, Húngaro, Iídiche, Italiano, Macedônio, Polonês, Români, Russo, Ruteno, Sérvio, Tártaro, Tcheco, Turco, Ucrâniano

Sérvia	2006	Sérvio	LRM: Albanês, Alemão, Bósnio, Búlgaro, Croata, Eslovaco, Húngaro, Români, Romeno, Ruteno, Tcheco, Ucrainiano
Suécia	2000	Sueco	LRM: Finlandês, Meänkieli, Lapão, Ídiche, Români
Suíça	1997	Alemão, Francês, Ieniche, Italiano, Romanche	
Ucrânia	2005	Ucraniano	LRM: Alemão, Bielorrusso, Búlgaro, Eslovaco, Gagaúze, Grego, Hebraico, Húngaro, Moldavo, Polonês, Romeno, Russo, Tártaro da Crimeia

Observação: ainda não ratificaram o ETS 148 os seguintes países: Albânia, Andorra, Azerbaijão, Bélgica, Bulgária, Estônia, França, Geórgia, Grécia, Irlanda, Islândia, Itália, Letônia, Lituânia, Macedônia, Malta, Moldávia, Mônaco, Portugal, Rússia, São Marinho e Turquia.

Fonte: <http://conventions.coe.int/Treaty/Commun/ChercheSig.asp?NT=148&CM=8&DF=&CL=FRE>.
Acesso em: 6 dez. 2014.

Observam-se situações diversas: (a) em Lichtenstein não há línguas minoritárias ou regionais, em vista da minúscula dimensão do país; (b) em Luxemburgo, também com território bastante reduzido, se reconhecem todas as línguas faladas dentro de suas fronteiras como oficiais, o mesmo acontecendo na Suíça; (c) na Noruega, as línguas regionais ou minoritárias são consideradas co-oficiais; (d) em Montenegro e Chipre se admite mais de uma língua como oficial, havendo ainda outras regionais ou minoritárias; (f) Espanha e Finlândia admitem línguas co-oficiais em regiões específicas, contando ainda com línguas minoritárias; (g) os demais países têm uma língua oficial e uma ou mais línguas regionais ou minoritárias.

Por outro lado, observe-se também como línguas oficiais em alguns países são regionais ou minoritárias em outros: o alemão, por exemplo, que é língua oficial na Alemanha, na Áustria, em Lichtenstein e em Luxemburgo, aparece como língua regional ou minoritária na Dinamarca, na Eslováquia, na Hungria, na Polônia, na República Tcheca, na Romênia e na Ucrânia. Merece destaque ainda o reconhecimento de línguas de comunidades específicas espalhadas por vários pontos da Europa, como o români, falado pelos ciganos e de origem índica, o qual integra o rol das línguas minoritárias da Alemanha, de Montenegro, dos Países Baixos, da Polônia, da República Tcheca, da Romênia e da Sérvia – o mesmo acontecendo com o ídiche ou o hebraico, usado por comunidades judaicas, os quais, em conjunto ou separadamente, são reconhecidos nos Países Baixos, na Polônia, na Romênia, na Suécia e na Ucrânia. Se recordarmos como ciganos e judeus europeus, ao longo da história e especialmente no século XX, sofreram uma série de perseguições, tendo sido submetidos a verdadeiro genocídio, o ato político de reconhecimento de suas línguas em países onde elas se falam representa por si só uma iniciativa de grande importância. Finalmente, não se pode deixar de observar que a Ucrânia, evitando a polêmica que você já conhece, reconhece como

línguas regionais ou minoritárias tanto o romeno quanto o moldavo!

Vamos considerar agora a situação do Brasil. Se você achou que alguns dos países citados, como a Polônia (com quinze) e a Romênia (com dezenove), têm muitas línguas regionais ou minoritárias, prepare-se: no Brasil, essa cifra é algo como que dez vezes maior e o elenco se apresenta muito mais variado. De um lado, temos as línguas ainda hoje faladas por descendentes dos imigrantes que aqui chegaram a partir da segunda metade do século XIX, a saber, alemão, árabe, chinês, coreano, espanhol, holandês, inglês, italiano, japonês, letão e pomerano (cf. BOLOGNINI; PAYER, 2005); de outro, há entre cento e trinta e cento e oitenta línguas indígenas espalhadas por todo o território nacional. Isso para não falar nas línguas ou falares de origem africana, em geral muito mescladas com o português e usadas em situações mais pontuais, como rituais religiosos. Em Minas Gerais, há os registros de Machado, em Milho Verde (*O negro e o garimpo em Minas Gerais*), e Queiroz, em Bom Despacho (*Pé preto no barro branco*) – havendo pelo menos o caso de uma, o calunga, falado na cidade de Patrocínio - MG, que parece ter uso mais amplo entre negros e brancos (cf. VOGT; FRY, 1996). No estado de São Paulo, também Vogt e Fry (1996) estudaram a língua do Cafundó. Trata-se, em todos os casos, de realidade que só há pouco tempo passou a ser efetivamente estudada, como, por exemplo, através do projeto “Enciclopédia de Línguas no Brasil”, da Universidade de Campinas (<http://www.labeurb.unicamp.br/elb>).

Para entender essa situação é preciso considerar a nossa história linguística. Durante os séculos XVI e XVII, sem dúvida o português não era a língua majoritária na então colônia, esse estatuto cabendo à chamada “língua geral” ou “tupi”, uma língua franca que já era utilizada por índios de diferentes etnias em toda a costa do Brasil, tendo sido adotada também pelos colonizadores portugueses. Tal era sua importância que dela se produziu uma primeira gramática, da autoria do Padre José de Anchieta, e no Colégio dos Meninos de Jesus, fundado pelos jesuítas em 1550, em Salvador (BA), ela constava como disciplina, ao lado do latim, sendo chamada, pelos padres, de “o grego da terra” (LEITE, 1938, p. 75). Segundo o testemunho de viajantes estrangeiros, a língua geral continuou a ser falada em São Paulo até fins do século XVIII, embora restrita apenas a uso doméstico (cf. HOLANDA, 1936, p. 94).

É apenas na segunda metade dos setecentos que a língua portuguesa se impõe, devido principalmente a dois fatores: (a) a descoberta do ouro no sertão, que faz com que um grande contingente de migrantes procedentes de espaços urbanos se transfira de Portugal para o Brasil; (b) a reforma do ensino levada a cabo pelo Marquês de Pombal, cuja primeira medida tinha em vista justamente as escolas primárias do Pará e Maranhão (medida de 3 de maio de 1757) e, em seguida, de todo o Brasil (17 de agosto de 1758), prescrevendo o seguinte: “será um dos principais cuidados dos Diretores estabelecer nas respectivas povoações o uso da língua portuguesa, não consentindo por modo algum que os meninos e meninas, que

pertencerem às escolas, e todos os índios, que forem capazes de instrução nesta matéria, usem da língua própria das suas nações ou da chamada geral” (*apud* CUNHA, 1985, p. 80).

O interessante, contudo, é que apenas a Constituição de 1988 declarou expressamente que “a língua portuguesa é o idioma oficial da República Federativa do Brasil” (art. 13). Anteriormente esse preceito permaneceu sem referência: nas constituições de 1824, logo após a Independência, e de 1891, que se segue à Proclamação da República, não há qualquer menção à língua nacional; a de 1934, no capítulo sobre o “Plano Nacional de Educação”, prescreve que o “ensino, nos estabelecimentos particulares, [será] ministrado no idioma pátrio, salvo o de línguas estrangeiras” (art. 150, parágrafo único, alínea “d”); na constituição de 1937 não se observa nenhuma referência à língua; na de 1946, são duas menções: uma sobre o ensino (“o ensino primário é obrigatório e será dado na língua nacional”, art. 168, inciso I) e a outra sobre os direitos políticos dos cidadãos, em que se declara que “não podem alistar-se a eleitores (...) os que não saibam exprimir-se na língua nacional” (art. 132, inciso II), tendo sido ambos os preceitos repetidos na constituição de 1967 (artigos 168 e 142, respectivamente). Finalmente, o texto de 1988 declara a língua portuguesa como “idioma oficial” da República, eliminando a restrição a que só podem votar os que o dominarem e dando a seguinte redação às normas relativas ao ensino: “o ensino regular será ministrado em língua portuguesa, assegurada às comunidades indígenas também a utilização de suas línguas maternas e processos próprios de aprendizagem” (art. 210, parágrafo 2º).

Esses passos apresentam alguns aspectos significativos. Antes de tudo, não declarar que o português é a “língua oficial” do Brasil não significa que não fosse, mas que isso estava subentendido – aliás, pelo próprio fato de que todas essas constituições sempre foram escritas em português. Mesmo antes da Independência, é certo que o português desempenhava o papel de língua oficial, já que era a única utilizada em documentos públicos no âmbito do Reino de Portugal e suas colônias. O fato de que, a partir de 1934, o “idioma nacional” comece a aparecer nas normas educacionais indica que alguma mudança havia ocorrido entre a última década do século XIX e a três primeiras do século XX, o que não é difícil de identificar, pois este é justamente o período da grande imigração europeia e oriental para o país (1887 a 1930). Para citar dados relativos apenas às comunidades mais numerosas, a quantidade de imigrantes italianos que aportou no Brasil é da ordem de 1,4 milhão, seus descendentes hoje, apenas na cidade de São Paulo, estando em torno de 5 milhões; em 1935, calcula-se que o número de falantes de alemão no Brasil era de cerca de 1,2 milhão, somando-se os imigrantes (cerca de 300 mil) e seus descendentes; o número de imigrantes japoneses, até a época da Segunda Guerra, é calculado em 190 mil, chegando seus descendentes hoje a 1,3 milhão (BOLOGNINI; PAYER, 2005). Ora, ao se estabelecerem, formando suas comunidades, esses migrantes naturalmente criavam escolas em que o ensino era ministrado nas línguas

maternas tanto de pais quanto de filhos. Assim se entende o motivo por que se introduziu, justamente no texto constitucional de 1934, a norma de que o ensino primário se faria apenas no “idioma nacional”, sendo principalmente por esse meio que as várias línguas acabariam esquecidas em favor do português e os imigrantes integrados na nacionalidade.

Mais interessante ainda é que nas disposições transitórias da Constituição de 1946 se preveja o seguinte: “O Governo nomeará Comissão de professores, escritores e jornalistas, que opine sobre a denominação do idioma nacional” (art. 35). Nesse caso, trata-se de dar solução a uma questão que se arrastava desde o século XIX, a de saber se o português do Brasil se chamaria “português” ou “brasileiro” (no estilo da polêmica envolvendo romeno e moldavo). A decisão, linguisticamente correta, é que a língua predominante no Brasil é a portuguesa e assim deve ser denominada. Que o assunto não é de pequena monta pode-se ainda constatar tendo em vista as prescrições, de 1946 e 1967, de que só podiam votar os cidadãos falantes de português, o que afetava principalmente as comunidades indígenas, mantidas em situação de tutela, sem reconhecimento efetivo de sua cidadania. É provavelmente por isso que a carta de 1988 cuidou de lembrar que o ensino, nessas comunidades, se poderá fazer “também” nas línguas próprias a cada uma delas – o advérbio garantindo que se trata de uma concessão que, todavia, expressa um reconhecimento nunca antes estabelecido em lei. Como a nova Lei de Diretrizes e Bases da Educação Nacional dedicou todo um capítulo à educação indígena, as línguas desses povos começam a ser retomadas e valorizadas, modificando a situação que perdurava desde as reformas pombalinas. Sem dúvida essa mudança de atitude foi o que permitiu que, no município de São Gabriel da Cachoeira (AM), em 2002, fossem legalmente reconhecidas como oficiais quatro línguas: o português, o nheengatu, o tukano e o baniwa, as três últimas faladas pela grande maioria da população (Lei Municipal 145, de 22 de novembro de 2002).

Trata-se de uma primeira experiência, ainda incipiente, mas que aponta no sentido da viabilidade de se adotarem políticas de incentivo à preservação das línguas regionais e minoritárias do Brasil, tendo em vista não só as indígenas, cuja conservação é sem dúvida prioritária, como também as raras africanas que ainda resistem e as das comunidades procedentes dos imigrantes europeus e orientais, esferas sobre as quais, aliás, a norma constitucional infelizmente nada prevê. Com efeito, a diversidade linguística constitui uma parte importante do patrimônio nacional, merecendo os mesmos cuidados que o patrimônio natural e material.

2 Línguas, dialetos, falares

Em termos estritamente linguísticos, Mattoso Câmara (s/d) assim define línguas, falares e dialetos:

- (a) Uma língua se distingue de outra pelo sistema de fonemas e pelo sistema de formas, bem como pelos padrões frasais em que essas formas se ordenam na comunicação linguística ou frase. Da estrutura específica de cada língua resulta a falta de inteligibilidade entre homens de línguas diversas, quando cada qual não aprendeu previamente o sistema de linguagem de cada um dos outros. (...) O que define uma língua, em face das demais, é a sua estrutura, que estabelece oposições específicas de fonemas e formas. De acordo com a estrutura, se tem uma nova língua a partir de um momento da evolução de uma língua dada (ex.: o português em face do latim) ou se distinguem num território contínuo duas ou mais línguas que são evolução de uma única língua (ex.: na península ibérica, a língua portuguesa, em fase da língua castelhana, ou espanhola, e da língua catalã, todas provenientes do latim). Há, entretanto, uma hierarquia nas oposições linguísticas e são as fundamentais, ou primárias, que definem essencialmente uma língua em face das demais. As oposições superficiais, ou secundárias, criam dentro de uma língua as divisões chamadas falares, que por sua vez são agrupáveis em dialetos. (p. 247-248).
- (b) Falares [são] línguas de pequenas regiões, através de um território linguístico dado, que se distinguem umas das outras por oposições superficiais dentro do sistema geral de oposições fundamentais que reúne todas numa língua comum. Os dialetos são a rigor conjunto de falares que concordam entre si por certos traços essenciais. Os falares caracterizam-se ainda, em face da língua comum, pela circunstância de pertencerem à língua cotidiana oral. (p. 175).
- (c) Do ponto de vista puramente linguístico, os dialetos são línguas regionais que apresentam entre si coincidência de traços linguísticos fundamentais. Cada dialeto não oferece, por sua vez, uma unidade absoluta em todo o território por que se estende, e pode dividir-se em subdialetos, quando há divergência apreciável de traços linguísticos secundários entre zonas desse território. A classificação dos dialetos e subdialetos de uma língua é, até certo ponto, convencional, pois depende dos traços linguísticos escolhidos para base de classificação; são sempre preferidos traços fonológicos e morfológicos porque a fonologia e a morfologia são aspectos de uma língua mais estáveis, mais sistemáticos e mais característicos de sua fisionomia. (p. 141).

Como você vê, as línguas podem ser distinguidas de diferentes perspectivas, não sendo tarefa simples estabelecer os critérios mais adequados. Em termos gerais, há duas possibilidades de estudo de uma língua: (a) como um sistema fechado, com enfoque no conjunto de sua gramática e de seu léxico, como nas descrições gramaticais e nos dicionários – o que Câmara define como “sistema de fonemas” (ou fonologia), “sistema de formas” (ou morfologia) e “padrões frasais” (sintaxe), a que é preciso acrescentar, sem dúvida, o léxico; (b) no uso concreto, algo em boa medida diferente, mais complexo e dependente de muitas variáveis não estritamente linguísticas, de ordem social, cultural e cognitiva. Para tornar mais claro como esses critérios são utilizados, distinguindo entre as oposições “fundamentais, ou primárias” e as “superficiais, ou secundárias”, tomaremos como exemplo o “sistema de fonemas”, estudado pela fonética e pela fonologia.

2.1 Fonética e fonologia

A fonética é a disciplina que estuda os sons emitidos pelo aparelho fonador humano e recebidos por seu aparelho auditivo, os quais podem ser medidos por equipamentos apropriados. Os sons, quando indicados foneticamente, são representados, entre colchetes, usando-se o alfabeto fonético internacional. Assim, a transcrição fonética das palavras portuguesas 'casa' e 'vaza' é, respectivamente,

[ˈkaza] / [ˈvaza],

não interessando que a norma ortográfica prescreva que uma se escreve com 's' e a outra com 'z'. Tomemos outro exemplo: 'chocar' e 'tocar' são foneticamente representadas como

[ʃoˈkar] / [toˈkar],

ou seja, o primeiro som de *chocar* (que, na escrita do português, é representado com dois símbolos gráficos, 'ch') corresponde ao que, no alfabeto fonético, é representado por [ʃ]; ao contrário, o primeiro som de *tocar* tem a mesma representação nos dois casos.¹⁶

A fonologia, por seu lado, é a disciplina que estuda os sons como são representados mentalmente e como se agrupam, numa determinada língua, de acordo com sua capacidade de atribuir distinções de significado às palavras. Considerando nosso segundo exemplo, em português [ʃoˈkar] e [toˈkar] se distinguem unicamente por um fonema, constituindo um par mínimo, ou seja, um par de palavras diferentes que têm tudo em comum menos um único som, o qual é suficiente para distinguir os respectivos significados. Isso indica que esses sons têm uma representação mental diferente e, por isso, são fonologicamente relevantes. Quando se trata de registrar a representação fonológica, colocamos os sons entre barras oblíquas:

par mínimo

/ʃoˈkar/ /toˈkar/.

A situação muda se comparamos, ainda em português, [t] e [tʃ]. O primeiro é o som inicial de 'teto'; o segundo, de 'tchau'. Em português, com exceção desta última palavra de origem estrangeira (grafada com três símbolos gráficos, 'tch'), o som [tʃ] não distingue significados. Tanto que uma mesma palavra, por exemplo 'tia', é pronunciada como [ˈtia] em algumas partes do Brasil e como [ˈtʃia] em outras, sem que se pense que as duas formas possam ter significados diferentes. Isso significa que a representação mental de [t] e [tʃ] é, em português, geralmente a mesma. Acusticamente percebemos a diferença, mas ela não muda o significado da palavra. Ao contrário, a diferença entre [p] e [b], acusticamente bem menor que entre [t] e [tʃ], é mentalmente mais significativa, tanto que [ˈpiko] e [ˈbiko], [ˈpasta] e [ˈbasta] são palavras diversas.

¹⁶ O acento, na representação fonética, é colocado antes da sílaba a que se refere.

Em resumo: o objeto da fonética são os *fonos*, um fone sendo um som concreto; o objeto da fonologia é o *fonema*, ou seja, a representação mental de um som que é capaz de distinguir significados. A fonologia estuda, por assim dizer, a dimensão cognitiva do som. De fato, apesar de todos os seres humanos terem as mesmas potencialidades fisiológicas para produzir e ouvir sons, sabemos que um brasileiro, um italiano ou um chinês não conseguem com a mesma facilidade produzir todos os sons nem decodificá-los quando os escutam. O que acontece é que tanto o aparelho fonador quanto o aparelho auditivo são comandados pela nossa cognição, sendo esta “moldada”, de diferentes maneiras, em diferentes lugares, a partir do nosso nascimento. A cognição de um brasileiro é moldada diferentemente da de um chinês: no primeiro caso, ela se molda de forma a diferenciar claramente os sons [r] e [l], o que não acontece com relação a um chinês. A cognição de um italiano é moldada para diferenciar claramente as consoantes intensas das simples, distinguindo, assim, entre

/'palla/ (bola) x /'pala/ (pá),

enquanto um estrangeiro geralmente não consegue perceber acusticamente nem executar foneticamente essa diferença. Portanto, se todos os sons, em princípio, podem ser igualmente produzidos e recebidos pelos aparelhos fonador e auditivo de todos os seres humanos (trata-se de uma capacidade da espécie), não são todos de fato realizados e percebidos com a mesma facilidade. Os falantes de uma língua produzem certos sons com mais facilidade que outros e percebem mais facilmente certas diferenças de sons. Um brasileiro não tem dificuldade em perceber a diferença entre

/'pãu/ (pão) x /'pau/ (pau),

mas, para um estrangeiro, as duas sequências podem parecer iguais, exatamente como para um brasileiro podem parecer iguais as sequências italianas

/'kasa/ (casa, 'casa') x /'kassa/ (cassa, 'caixa'),

ou as sequências inglesas

/'tin/ (tin, 'lata') x /'θin/ (thin, 'magro', 'fino').

A percepção obviamente se reflete na realização: o que é difícil de ser percebido é também difícil de ser realizado. Isso decorre do fato de que a decisão do que constitui um *fonema* se faz de maneira diferente em cada língua. Um som que é fonema em uma língua pode não ser em outra. Assim, enquanto o português considera [t] e [tʃ] como duas realizações possíveis do mesmo fonema /t/, o italiano admite que se trata de dois fonemas diferentes e diferencia os pares mínimos

/'tindʒere/ (tingere, 'tingir') x /'tʃindʒere/ (cingere, 'cercar').

Enquanto em português ou em italiano a alternância entre [t] por [θ] não muda o significado de nenhuma palavra – no máximo, podemos achar esquisita a pronúncia da pessoa –, em inglês o contraste é significativo, como em

/tri/ (*tree*, 'árvore') x θri (*three*, 'três').

Desse modo, ainda que os fones que o ser humano pode produzir sejam muitos, os fonemas de cada língua são limitados – por volta de 30 em boa parte delas – e os fonemas são distinguidos autonomamente em cada comunidade de falantes, ou seja em cada língua. Essa é uma das razões que torna difícil o aprendizado de uma língua estrangeira quando o cérebro de um falante está já moldado e, portanto, acostumado a agrupar cognitivamente os fones em fonemas. As distinções que são importantes na língua que queremos aprender, mas não na que falamos, são as mais difíceis. Assim, é fácil ouvir um brasileiro ou um italiano pronunciarem a palavra *think* (pensar) do inglês como se fosse *tink* (como em *tinker*, funileiro) ou *sink* (afundar). Isso porque tanto o português quanto o italiano possuem os fonemas /t/ e /s/, mas não possuem /θ/.

Contudo, observe-se que não pronunciamos o mesmo fonema usando sempre o mesmo fone, o que depende de vários fatores. Dentre os mais importantes estão o fator diatópico, ou seja, geográfico, e o contextual, isto é, o contexto criado pelo som que aparece imediatamente antes ou imediatamente depois daquele que pronunciamos. Nestes casos, trata-se de variações que não se incluem nas que Câmara define como “fundamentais ou primárias”, as quais distinguem as línguas umas das outras, mas das variações “superficiais ou secundárias”, que “criam dentro de uma língua as divisões chamadas falares, que por sua vez são agrupáveis em dialetos”.

Assim, por exemplo, se em Belo Horizonte se pronuncia [ˈtʃia], em outros lugares do Brasil se diz [ˈtia], o que configura uma diferença diatópica do português. Um outro exemplo é a pronúncia do [s] final no Rio de Janeiro, realizado como se fosse um [ç], muito parecido ou até igual ao primeiro som da palavra *chocar* /çɔˈkar/. Ora, se em ‘chocar’ [ç] é um fonema (tanto que substituindo-o por [s] mudamos o significado: chocar/socar), na pronúncia carioca de ‘amigos’ como [amiguç], o [ç] final é apenas uma realização diferente do fonema /s/, ou seja, um *alofone*, quando dois ou mais fones são admitidos na realização de um mesmo fonema. Assim, tanto o [ç] final do Rio de Janeiro é não mais que um alofone de /s/, quanto o [tç] mineiro é um alofone de /t/.

Observe, todavia, que não é em qualquer situação que /t/ admite a realização como /tç/ em Belo Horizonte, mas apenas quando seguido do fonema /i/ (escrito como ‘i’ ou como ‘e’), ou seja, a possibilidade do alofone depende do contexto em que ocorre o fonema, o que se define como fator contextual: assim, a pronúncia [ˈleitçi] (*leite*), em português, é possível como uma variante de /ˈleiti/, mas /ˈleitu/ (*leito*) não admite um alofone em [tçu]. A influência do contexto é clara: como /i/ é uma vogal palatal, ela palataliza o fonema /t/, possibilitando que seja realizado como /tç/ – o que se chama assimilação regressiva, já que o fone posterior, [i], produz efeito no anterior, [t]. Do mesmo modo, o fonema /s/ é pronunciado como [ç] no

Rio de Janeiro e em outras regiões somente em alguns contextos: basicamente, em posição final e antes de consoante, como em ‘mapas’ ,[‘mapaç], e ‘mascar’, [maç‘kar], mas não em ‘sapo’, [‘sapu], e ‘massa’, [‘masa].

Também o mesmo fator situacional se observa na tendência de aparecimento – o que se chama epêntese – de uma semivogal [i] em palavras oxítonas terminadas em [as] e [es], própria do português do Rio de Janeiro, dentre outras regiões: [‘mas] pronunciado como [‘maiç] e [‘ves] como [‘veiç]. O quanto isso é importante pode ser observado em registros de linguagem sensíveis, como o da poesia. Assim, por exemplo, uma canção de carnaval carioca dizia:

Existem quatrocentas mil mulheres a mais
Da Penha do Posto Seis,
São mais de dez mulheres pra cada rapaz
Só eu não tenho vez –

as rimas só acontecendo caso ‘rapaz’ se pronuncie como [ra‘paiç], para rimar com [maiç], e ‘vez’ como [‘veiç], rimando com [‘seiç].

2.2 Morfologia

O que distingue as línguas em termos morfológicos deve ser considerado também a partir da mesma lógica: interessam sempre os traços significativos, capazes de constituir um sistema de categorias gramaticais. Como na fonética e na fonologia podemos analisar os fones, que são realizações concretas, e os fonemas, que são representações mentais e distinguem significados, assim também na morfologia podemos distinguir entre os morfos, que são realizações concretas, e os morfemas, que são as menores unidades de significado.

Tomemos alguns exemplos. Em português, a palavra ‘meninos’ pode ser segmentada em três morfos, cada um expressando um significado: {menin} expressa o significado lexical; {o} expressa o significado de gênero masculino; {s} expressa o significado de número plural. Neste caso, cada forma, ou seja cada morfo, expressa um único significado. Mas em línguas como o português e, em geral, nas línguas indo-europeias (chamadas de flexivas ou fusivas), isso não acontece com frequência. O mais comum é que haja mais de um significado, ou seja, mais de um morfema, para um mesmo morfo. Comparando a palavra ‘meninos’ com seu equivalente em italiano, *bambini*, constatamos o seguinte: há o morfo que expressa o morfema lexical {bambin} seguido por um único morfo que expressa dois morfemas, {i}, pois vale tanto para o significado de gênero {masculino}, quanto para o significado de número {plural} – o plural feminino, *bambine*, sendo assim analisável: {bambin}, conteúdo lexical + {e}, gênero feminino e plural.

Essa sobrecarga de morfemas num único morfo é evidente na morfologia do verbo em

português. O imperfeito do verbo *amar*, '(eu) amava' pode ser segmentado em {am}, morfema lexical, como em 'amor', 'amante' etc, seguido por {av} e {a}, os quais expressam mais de um significado: (a) {av} expressa pelo menos os significados de modo indicativo (assim 'amava' contrapõe-se, por exemplo, a 'amasse', imperfeito do subjuntivo), tempo passado ('amava' está em oposição a 'amo' e 'amarei', presente e futuro, respectivamente), e aspecto durativo ('amava' distingue-se de 'amei', que também é uma forma do passado, por expressar duração); (b) já o morfo {a} expressa os morfemas de primeira pessoa (em contraposição com 'amavas') e de número singular (opondo-se a 'amávamos'). Um exemplo extremo é o caso da palavra 'é', do verbo 'ser': com um único fonema e, portanto, um único morfo, indicamos uma série de morfemas, ou seja, de significados, entre os quais pelo menos os seguintes: morfema lexical, morfema de modo indicativo, morfema de aspecto durativo, morfema de tempo presente, morfema de terceira pessoa, morfema de número singular.

Todavia, o contrário também pode acontecer, ou seja, que um mesmo morfema, um mesmo significado, seja representado por mais de um morfo. De fato, em português, para o significado {plural} nem sempre o morfo seria {s}, porque o plural de 'mar' é 'mares', o de 'caracol', 'caracois', o de 'caminhão', 'caminhões'. Como você vê, tanto {s}, quanto {es}, {is}, {ões} e outros são morfemas do mesmo morfema {plural}. Quando vários morfemas são realizações concretas do mesmo morfema, são chamados de *alomorfos*.

A morfologia se divide em morfologia gramatical e morfologia derivacional. A morfologia gramatical serve para fornecer significados gramaticais ao mesmo lexema, ou seja, ao mesmo significado lexical da palavra. Por exemplo 'menino' e 'menina' são duas formas gramaticais do mesmo lexema. A única distinção está no gênero. Assim todas as formas do verbo 'amar' (*amo, amamos, amei, amaria* etc.) são formas diferentes do mesmo lexema. Ao contrário, a morfologia derivacional trata da formação de palavras, isto é, lexemas novos a partir de um lexema que serve como base. Por exemplo, não podemos dizer que 'comum', que é um adjetivo, e 'comunidade', que é um nome, são o mesmo lexema. O segundo, contudo, é derivado do primeiro através de um processo de sufixação, ou seja, pelo acréscimo ao final da palavra de um morfema (*-idade*) que tem a capacidade de fazer com que um adjetivo se torne um nome e um novo lexema. Com o tempo pode até acontecer que os falantes percam a percepção que uma palavra é derivada de outra. Ninguém mais percebe que 'amigo' é derivado da mesma base de 'amar'. A estratégia da derivação é extremamente comum nas línguas como forma de ampliar o vocabulário. É uma estratégia muito transparente, porque o sufixo ou o prefixo possui um significado imediatamente interpretável. Por exemplo, o contrário de 'jovem' é 'velho', sendo necessário que saibamos o significado de ambas as palavras, porque não há nada que nos diga que uma é o contrário da outra. Mas com relação a 'feliz' e 'infeliz', é suficiente saber o que significa 'feliz' e que o

prefixo *in-* significa negação para entender o significado do segundo termo.

A derivação geralmente se faz de três maneiras:

1. Através de um afixo, ou seja, acrescentando-se um morfema antes (prefixo), depois (sufixo) ou, em certas línguas, no interior (infixo) da raiz da palavra;
2. Através da composição, isto é, agrupando duas palavras, como em ‘guarda-chuva’, termo composto de um verbo e de um nome;
3. Através do processo chamado de *conversão*, um mecanismo muito comum nas línguas com morfologia muito simples, como o inglês, que consiste em usar uma mesma forma com funções morfológicas diferentes, constituindo, de fato, lexemas diferentes: em português, por exemplo, ‘poder’ pode ser verbo mas pode ser também nome, como na frase ‘o poder do ministro é muito grande’; ‘rápido’ pode ser adjetivo mas pode também ser advérbio, como em ‘andar rápido’; em inglês, *back* e *round* podem ser:

QUADRO 2

Exemplos de conversão em inglês

Classe	<i>back</i>	<i>round</i>
Substantivo	<i>I have a pain in my back</i> , ‘estou com dor nas costas’	<i>rounds of paper</i> , ‘círculos de papel’
Adjetivo	<i>the back door</i> , ‘a porta dos fundos’	<i>a round table</i> , ‘uma mesa redonda’
Verbo	<i>please, back the car</i> , ‘por favor, chegue o carro para trás’	<i>to round a figure</i> , ‘arredondar uma quantia’
Advérbio	<i>look back!</i> , ‘olhe para atrás’	<i>the earth goes round</i> , ‘a terra gira’
Preposição	<i>I go back home</i> , ‘eu vou de volta para casa’	<i>to travel round the world</i> , ‘viajar ao redor do mundo’

Muito frequente em inglês é a conversão verbo-nome ou adjetivo: *to look* (“olhar”) e *give a look* (“dar uma olhada”); *to nail* (“pregar”) e *give me a nail* (“dê-me um prego”); *to good* (“abonar”) e *a good man* (“um homem bom”); *to slow* (“diminuir a velocidade”) e *a slow student* (“um estudante burro”); etc.

3 O diassistema

Como você vê, se uma língua se define por sua fonética, morfologia, sintaxe e léxico, nem tudo é tão simples, uma vez que se admitem variáveis em maior ou menor grau, como os alofones e os alomorfos. A possibilidade da existência de variantes atinge todos os níveis gramaticais, ou seja, nenhuma língua constitui um bloco absolutamente homogêneo, sua riqueza e vitalidade estando expressas justamente pela capacidade de criação e variação

constantes.

Tomemos mais um exemplo de nossa própria língua. O fato de que a conjugação verbal se apresente de forma variada, tanto nos registros escritos quanto orais, no Brasil e em Portugal, não faz com que se trate de línguas diferentes, mas apenas de diversas variantes. No quadro seguinte, você encontrará um rol de possibilidades de conjugação do verbo 'ir', nenhuma delas agramatical, embora algumas formas não sejam admitidas pela gramática normativa:

QUADRO 3

Tipos de conjugação verbal no português contemporâneo

Tipo 1	Tipo 2	Tipo 3	Tipo 4	Tipo 5	Tipo 6	Tipo 7
eu vou	eu vou	eu vou	eu vou	eu vou	eu vou	eu vou
tu vais	tu vais	tu vai	você vai	você vai	você vai	você vai
ele vai	ele vai	ele vai	ele vai	ele vai	ele vai	ele vai
nós vamos	nós vamos	nós vamos	nós vamos	a gente vai	nós vai	nós vai
vós ides	vocês vão	vocês vão	vocês vão	vocês vão	vocês vão	vocês vai
eles vão	eles vão	eles vão	eles vão	eles vão	eles vão	eles vai

Antes de avaliar os sete tipos, convém esclarecer o que se considera linguisticamente como *gramatical* ou *agramatical*. Não se trata de definir se algum uso é "errado" do ponto de vista da gramática normativa, mas sim se está de acordo com a gramática da língua internalizada pelos falantes ou a contrária. Assim, 'nós vai' e 'eles vai', embora inadequados ao registro formal do português, não oferecem desvios de compreensão, sendo, portanto, gramaticais. Já '*eu vamos' ou '*você vão' são formulações agramaticais e sem sentido.¹⁷

Observe que na sequência dos tipos o traço principal é a diminuição das formas verbais: o primeiro tem seis, o último, duas. Note ainda:

- (a) No primeiro tipo, pertencente ao nível mais formal, falado e escrito, os pronomes são mesmo dispensáveis, já que o verbo, em cada pessoa do singular e do plural, apresenta uma forma própria – ou seja, seis.
- (b) No segundo, uma forma verbal difícil e com uso cada vez mais restrito, mesmo na linguagem formal, a segunda pessoa do plural 'ides', é substituída pela terceira, com a substituição do pronome 'vós' por 'vocês', o que se observa tanto em Portugal quanto no Brasil, embora mais destacadamente no último. Note que os pronomes 'você' e 'vocês' têm o sentido de segunda pessoa (com quem se fala), mas concordam com o verbo em terceira pessoa, uma vez que sua origem é uma forma arcaica de tratamento respeitoso, 'Vossa Mercê' – equivalente, portanto, a fórmulas

¹⁷ As formas agramaticais são sempre marcadas por um asterisco, o que implica que não são admissíveis pela gramática de uma determinada língua.

ainda em uso como 'o(s) senhor(es)', 'a(s) senhora(s)', 'Vossa(s) Excelência(s)', 'Vossa(s) Magnificência(s)' etc, todas de segunda pessoa, mas exigindo que o verbo se ponha na terceira pessoa do singular ou do plural. A norma coloquial culta em Portugal e em algumas regiões do Brasil, especialmente no Nordeste, adota preferencialmente este tipo, com cinco formas do verbo.

- (c) No terceiro tipo, usa-se 'tu' como pronome de segunda pessoa, mas o verbo tem a mesma forma que a terceira pessoa, 'vai'. Embora a gramática normativa considere que se trata de uma opção incorreta, ela tem largo uso em regiões como o Rio de Janeiro e o Sul do Brasil, apresentando o verbo, como se vê, quatro formas.
- (d) O quarto tipo faz a substituição de 'tu' por 'você', ficando o verbo com quatro formas. Essa é uma alternativa corrente no português padrão do Brasil, tanto escrito, quanto falado, a qual não fere nenhuma regra da gramática normativa.
- (e) Entre os tipos 4 e 5 a única diferença está na alternância do pronome de terceira pessoa 'nós' com 'a gente', exigindo-se, no segundo caso, a concordância com o verbo em terceira pessoa. Assim, de quatro formas (vou/vai/vamos/vão), passa-se a três (vou/vai/vão), também sem ferir nenhum preceito da gramática normativa. Pode-se dizer que o português coloquial culto do Brasil alterna entre esses dois usos.
- (f) No sexto tipo, encontramos o pronome 'nós' com o verbo em terceira pessoa, um uso não admitido pela gramática normativa, mas difundido na linguagem coloquial não-culta, conservando o verbo as três formas do tipo 5 (vou/vai/vão).
- (g) Finalmente, no último tipo, um registro bastante popular e localizado de linguagem, o verbo reduziu-se a duas formas (vou/vai).

É importante observar que quanto mais se reduz a marcação morfológica das pessoas verbais, tanto mais o uso dos pronomes se torna obrigatório. Isso quer dizer que, tipologicamente, de acordo com a classificação proposta por Chomsky (*Lectures on Government and Binding*), se a primeira coluna apresenta uma situação típica de uma língua *pro-drop* (do inglês *pronoun-dropping*, elisão do pronome), como são o latim, o italiano, o espanhol, o grego, o turco, o basco etc, quanto mais se avança na sequência mais se configura uma língua não *pro-drop*, ou seja, do tipo do inglês e do francês. Isso significa que a marcação da pessoa verbal deixa de ser feita à esquerda, através de sufixos, passando a ser feita à direita, por meio dos pronomes. Consequência disso é que os próprios pronomes passam por um processo que se chama de *gramaticalização*, ou seja, perdem seu conteúdo lexical passando a ter um significado gramatical, o que geralmente implica também em perda de material fonético. Embora não tenhamos registrado no nosso quadro as variações dos pronomes, 'você vai', 'vocês vai' e 'eles vai' realizam-se, efetivamente, como /se'vai/, /ses'vai/ e /es'vai/.

Os fatos acima, relativos ao português, permitem que consideremos a língua como um *diassistema*. Normalmente, como vimos, se diz que uma língua é um sistema de traços fonéticos, morfológicos, sintáticos etc. Na verdade, as chamadas “línguas de cultura”, ou seja, as línguas de sociedades complexas, como o inglês, o português, o italiano, o chinês e muitas outras, mas com certeza uma minoria das línguas do mundo, não devem ser consideradas somente um sistema, mas um diassistema, ou seja um sistema de sistemas.

Os vários sistemas que constituem essas línguas são consequência do cruzamento de algumas variáveis de natureza sociolinguística, tais como:

(a) Variável diacrônica – a língua varia através do tempo. Continuando a ter o português como exemplo, é claro que como ele se fala hoje difere de como se falava há um século atrás. Qualquer língua muda com o tempo, acrescentando novos elementos e eliminando outros em qualquer nível, seja fonético, morfológico, lexical etc.

(b) Variável diatópica – a língua varia através do espaço. De maneira mais ou menos evidente, todas as línguas variam de um lugar para o outro. O português do Rio de Janeiro, como vimos, é caracterizado por vários traços que o diferenciam do de Belo Horizonte ou de São Paulo.

(c) Variável diastrática – a língua varia através dos grupos sociais. Um indivíduo pertencente ao grupo masculino fala de maneira mais ou menos diferente (a intensidade da diferença dependendo da língua) de um indivíduo que pertença ao grupo feminino; um jovem fala de maneira diferente de um adulto, de uma criança ou de um velho; mas, principalmente, essa variação se dá em virtude dos processos de escolarização.

(d) Variável diafásica – a língua varia através das situações comunicativas. Quando se muda a situação, usamos uma língua diferente. Por exemplo, não usamos a mesma língua se estamos conversando com amigos num bar, apresentando um projeto para um potencial comprador de nossos serviços, brigando com a namorada, discutindo durante uma reunião profissional etc. Também não usamos a mesma língua ao escrever um e-mail contando um acidente de automóvel para um amigo ou expondo o mesmo acidente para a nossa companhia de seguro – e muito menos se estamos escrevendo nossa tese de doutorado, um romance etc. Para cada situação existe uma linguagem apropriada – e várias outras possibilidades, que o sistema oferece, mas que não são apropriadas para aquela situação.

(e) Variável diamésica – a língua varia dependendo do meio usado para transmiti-la. Quando falamos, somos condicionados pelo meio (as ondas sonoras) e pelo canal (o ar), assim como quando escrevemos somos condicionados pelo suporte e pelas características da escrita. O mesmo acontece com relação a outros meios, como o

cinema, o rádio, a televisão, a *internet* etc. Ninguém escreve da mesma maneira que fala ou usa a mesma linguagem num chat, num e-mail, ou numa obra literária.

A figura 1, abaixo, mostra algumas das relações entre essas variáveis:

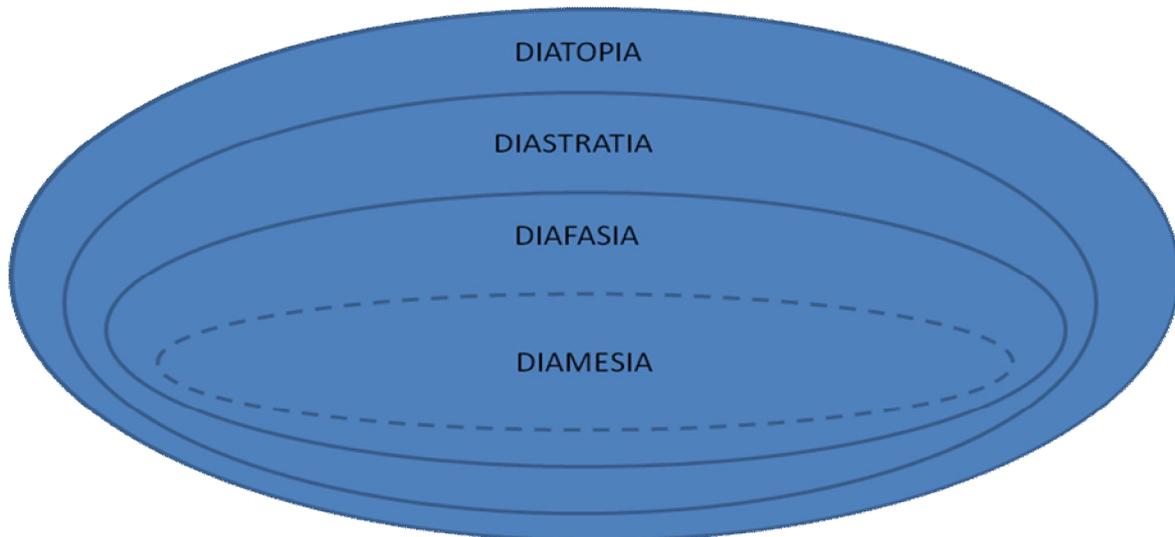


FIGURA 1 - O diassistema

Como se vê, não representamos a diacronia, porque as relações entre as variáveis podem ser analisadas somente num sistema sincrônico, ou seja, dado o diassistema de uma certa época, ele gerará vários sistemas. Noutros termos: cada época apresenta seu próprio diassistema.

Note como há, na figura, vários círculos ovais: o círculo da diatopia contém o da diastratia, que contém o da diafasia, que, finalmente, contém o da diamesia. Repare ainda que este último, contrariamente aos outros, não é desenhado com uma linha contínua, pela razão que logo veremos. O fato de a diatopia ser o círculo mais externo e de a progressão se dar a partir dele, passando pela diastratia e depois pela diafasia, até atingir a diamesia, tem um sentido, pois representa o percurso seguinte:

1. Primeiramente, aprendemos a falar a língua da diatopia, não somente enquanto a língua do lugar onde aprendemos a falar (o português, por exemplo), mas a variedade de língua que se fala nesse lugar, com suas características fonéticas, lexicais e morfossintáticas (o “mineirês”, por exemplo). Quem nasce e aprende a falar em Belo Horizonte aprende, portanto, uma variedade diferente de quem nasce no Rio de Janeiro. O condicionamento diatópico é o primeiro que recebemos e

ninguém pode evitá-lo – apenas os locutores profissionais, depois de muito esforço, aprendem a dissimulá-lo.

2. Em um segundo momento, somos condicionados pelo grupo de que fazemos parte: se somos homens, seremos condicionados pela maneira masculina de falar; se crescemos em um ambiente familiar culto, aprenderemos a nos expressar de maneira mais parecida à de nosso ambiente; se estudamos, aprenderemos a nos expressar de maneira mais complexa; e podemos ser condicionados de várias outras maneiras, dependendo do(s) grupo(s) em que estamos inseridos.
3. Em terceiro lugar, somos condicionados pelas situações de que temos experiência suficiente. Todos têm experiências de situações de intimidade ou de informalidade, mas nem todos terão de situações profissionais altamente complexas – e, com certeza, quem as tiver, terá somente de algumas, não de todas. Poucos teremos experiência de discursos públicos e pouquíssimos de nós terão de muitas situações públicas diferentes etc.
4. Somente como último estágio alcançamos a variedade diamésica, quando a alcançamos. De fato, se todo mundo aprende a falar, nem todo mundo aprende a escrever e menos gente ainda aprende a comunicar-se através de outros meios, como a televisão, o rádio, a *internet*. Por isso a linha do círculo mais interno não é contínua, porque nem sempre é alcançada. Em suma, nem todos experimentam variações diamésicas, mas somente uma única diamesia, a fala.

Naturalmente, cada pessoa possui uma capacidade diferente de lidar com as variações para criar diferentes sistemas a serem usados dependendo das exigências comunicativas. Esse conjunto de sistemas é chamado de repertório de cada falante. Ninguém possui todos os sistemas, mas algumas pessoas possuem mais sistemas do que outras. Quem teve acesso à escolarização e experimentou durante a própria formação linguística um número mais variado e amplo de situações possui claramente mais sistemas do que quem não teve essas oportunidades. Mas alguns sistemas de baixa complexidade, como, por exemplo, o caipira, não são adquiridos através da instrução ou da experiência situacional em sociedades complexas. Portanto, um caipira pode não ter acesso a nenhuma outra variedade do repertório, mas isso não significa que a sua variedade seja incluída no repertório de quem possui uma variedade mais ampla.

A figura 2, que segue, mostra algumas possíveis variedades do português brasileiro:

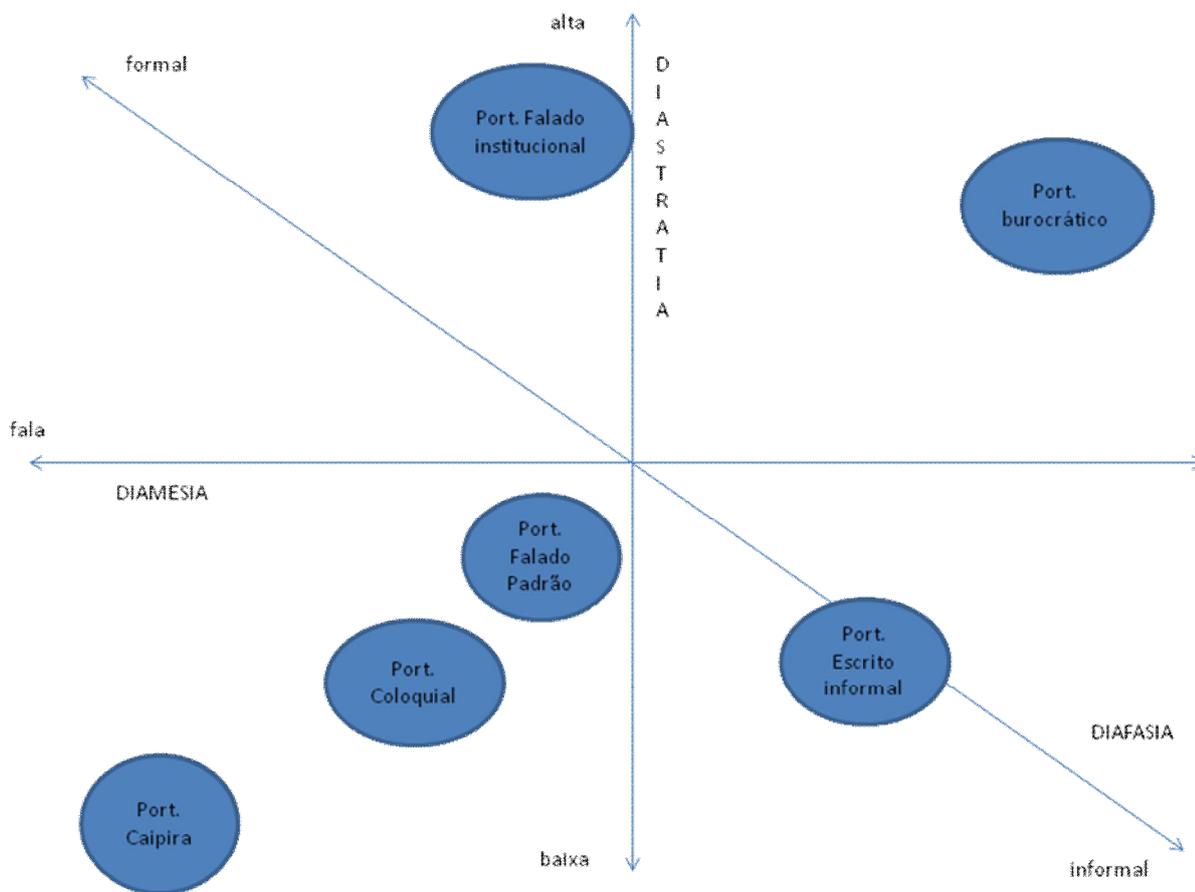


FIGURA 2 - Algumas variedades do português brasileiro

São necessários alguns comentários. Em primeiro lugar, o gráfico deve ser imaginado como tridimensional: a linha horizontal representa a largura do espaço, a linha vertical, sua altura, e a linha oblíqua a profundidade, que obviamente não pode ser adequadamente representada em uma superfície. Dito isso, a linha vertical representa a diastratia, do ponto mais baixo até o mais alto; a linha horizontal representa a diamesia, com o polo da oralidade mais interativa e não planejada à esquerda e o da escrita mais planejada à direita; a linha oblíqua representa a diafasia, com as situações mais informais na parte baixa à direita e as situações mais formais na parte alta à esquerda. Falta, além da variação diacrônica, também a diatópica. Isso se deve ao fato que este gráfico se refere a uma diatopia, pois cada diatopia possui um conjunto de variações diastráticas, diafásicas e diamésicas próprias. O gráfico deveria, portanto, ser repetido para cada diatopia.

O que se mostra é que, no espaço tridimensional de um diassistema linguístico, é possível identificar a posição ocupada por cada sistema com relação aos demais. Os seis sistemas indicados para o português do Brasil não são claramente os únicos do diassistema brasileiro, servindo unicamente para exemplificar a relação entre sistemas e variáveis. Não

existe um número definido de sistemas, porque o espaço linguístico deve ser considerado como um *continuum*, inteiramente ocupado por inúmeras possibilidades. Se compararmos sistemas distantes, as diferenças são claras, mas, entre um e outro, existem muitas realizações possíveis.

O que diferencia um sistema do outro dentro de do mesmo diassistema? Obviamente estamos sempre dentro da mesma língua. Portanto, muitos traços, como preposições, artigos, várias conjunções e muitos lexemas são iguais em todos os sistemas. Há outros traços que podem participar de mais de um sistema, mas geralmente com frequência bem diferente. Outros ainda são específicos de um sistema ou de outro.

Para diferenciar os sistemas devemos considerar pelo menos o seguinte:

1. Traços que podem estar presentes em todos os sistemas, mas cuja frequência varia de um sistema para o outro;
2. Traços específicos de um sistema, como léxico especializado, alguns usos morfossintáticos etc;
3. Coocorrência de traços, uma vez que é muito frequente que alguns sistemas sejam caracterizados não pela presença de um ou outro traço específico, mas pelo fato de que certos traços ocorrem em conjunto somente naquele sistema, ou seja, tomados individualmente, esses traços encontram-se também em outros sistemas, mas somente em um sistema eles ocorrem juntos.

Quando se afirma que somente as línguas de sociedades complexas são diassistemas, entende-se que o diassistema se produz quando uma língua necessita de muitos sistemas porque com ela precisamos agir de maneira diferente, em muitas circunstâncias diferentes, para fazer várias coisas de maneira diferente. A língua de uma pequena comunidade que não possua escrita e, portanto, não conheça variação diamésica, não possua estratificação social e, portanto, não experimente variação diastrática, e tenha necessidade de lidar com uma variedade de situações limitada (pois não possui escolas, hospitais, burocracia etc.), essa língua não é considerada um diassistema, mas um sistema simples.

4 O que é uma língua

Como você viu, uma língua é algo extremamente complexo, em que interfere um número bastante elevado de variáveis significativas, de ordem externa e interna, as quais são estudadas por diferentes ramos da linguística. Quando perguntamos o que é a língua portuguesa, por exemplo, é necessário considerar seus diversos sistemas em termos diacrônicos, com suas diatopias, no interior das quais se encontram as respectivas diastratias, diafalias e diamesias.

Assim, se formas gramaticais como a mesóclise, a ênclise e o mais-que-perfeito simples se pode dizer que não integram mais o português coloquial contemporâneo do Brasil, isso não implica que não sejam categorias da língua portuguesa, mesmo da brasileira contemporânea, uma vez que se mantêm, por exemplo, na língua escrita.

Do mesmo modo, ainda que não usemos, na comunicação coloquial, verbos na segunda pessoa do singular, podemos nos valer deles ao escrever um poema, o que é muito comum (“*Embebedaste* minha vida de vinhos ora suaves, ora secos/ *Me deste* a melhor comida e às vezes a melhor cachaça...”, Casa Nova, 2004, p. 31) – e até a forma menos usual dentre todas, a segunda pessoa do plural, não deixa de se manter corrente, por exemplo, na linguagem religiosa, que costuma ser bastante conservadora (“Pai nosso que *estais* no céu, santificado seja o *vosso* nome, (...) não nos *deixeis* cair em tentação, mas *livrai*-nos do mal”). Aliás, a língua de uma cultura complexa expõe seus usuários desde cedo a um conjunto amplo de variáveis, seja em situações mais informais, como nas canções infantis (“O anel que *tu me deste* / era vidro e seu quebrou, / o amor que *tu me tinhas* / era pouco e se acabou.”), seja através do aprendizado escolar. Apenas como último exemplo, registre-se que o sambarenredo da Imperatriz Leopoldinense, no carnaval de 2010, simplesmente cantava, na segunda do plural(!): “Oh Deus pai / *Iluminai* o novo dia, / *Guiai* o divino destino/ Peregrino em harmonia...”

* * *

LEITURA COMPLEMENTAR

Leia o texto abaixo e procure descobrir em que língua está escrito, muito semelhante, como você constatará, ao português:

Esta eideia de qu'uas lhénguas son melhores qu'outras chegou a tener muitos defensores, que magicórun¹⁸ theories mais ou menos mirabulantes. Dante, por eisemplo, dezie que las lhénguas bulgares, nun serbien para falar de cousas profundas. Umberto Eco fala-mos nun tal baron de Ryckholt, un flamengo que defendie que la lhéngua flamenga era mais antiga, falada deçde l bércio de l'houmanidade. Todas las outras éran dialetos defíciles d'antender. L sueco Andreas Kempe, dezie que Dius, an ne Paraíso, falou an sueco, Adan an dinamarqués i, esta ye la melhor, la serpiente qu'ancantou Eba falaba francés. Eiqui na bezina Spanha fazírun scola las palabras de l Amperador Carlos I para quien l castelhana era la lhéngua mais apropiada para falar cun Dius.

Fui nesta rateira nacionalista que se deixórun caer tamien alguns filózofos i scritores de ls mais afamados. L romántico Herder, por eisemplo, dezie que la lhéngua ye l melhor speilho de l'alma nacional. L filózofo Leibniz, defendie que l aleman ye la lhéngua que cunserbou la cara mais natural ou adânico. Ou seia, aqueilha que stá mais acerca de Dius i, bien atendido, la lhéngua mais porfeita que las outras.

Mas todas estas theories dében ser atendidas no sou spácio i no sou tiempo. Alguas pertécen a ua era an que las nações ouropeias buscában, por todos ls meios, las razones para defenir las sues frunteiras i para lhegitimar la sue eijistêcia. Las lhénguas fúrun, desde siempre, l stendarte de ls pobos dominados. Fúrun eilhas que criórun las gientes, que fabricórun ls pobos i ajudórun a formar las nações. Las lhénguas son l arado i la charrua qu'abrírun ls sucos donde stan sembradas las sementes de l'houmanidade. Hoije ls tiempos son outros. Mas las lhénguas – todas las lhénguas – stan ende para mos dar cunta que somos hardeiros desta bariedade lhenguística. Ye ua hardança i ua riqueza que tenemos l'oubrigaçon de cunserbar porque eilha representa aquilho que de mais sublime tenemos cumo espécie.

* * *

É provável que você não tenha podido identificar a língua em que o texto se encontra escrito (se conseguiu, parabéns!, é provável que você seja um de seus quinze mil falantes). Trata-se do mirandês, por alguns considerado um dialeto do asturiano, falado no concelho de Miranda do Douro, nordeste de Portugal, país em que é reconhecido, ao lado do português, como língua oficial.

¹⁸ òImaginarämö.

Até fins do século XX, restringiu-se à oralidade. Em 1882, o filólogo português Leite de Vasconcelos publicou estudos sobre ela, recolhendo ainda contos, adivinhas, provérbios, fábulas e canções em mirandês. Foi seguido por outros, como Bernardo Fernandes Monteiro, que traduziu para o mirandês os quatro evangelhos, trabalho publicado apenas parcialmente, a partir de 1896. Dos anos 1960 para cá, cresce a produção nessa língua, que, atualmente, como outras línguas regionais e minoritárias da Comunidade Europeia, busca manter-se, firmando-se também como língua escrita. Em 2008 sua norma ortográfica foi estabelecida pela Câmara Municipal de Miranda do Douro.

O texto a seguir, de que o anterior é um extrato, reflete bem o desejo de reconhecimento da parte de seus falantes:

Para quei sirbe l mirandês?

António Bárbolo Alves

La lhéngua mirandesa – cumo muitas outras ditas minoritárias, chamadas dialetos ou an situaçon dialetal – salidas d'un mundo meio zbarrulhado ou an que yá nun queda piedra subre piedra, cumo l chamado mundo rural ou tradicional, lhieban a las cuostas ua cruç mui pesada. Ye l fardo de, segundo alguns, nun serbíren para falar de l mundo moderno. Ora ua lhéngua que solo serbisse para falar de bacas (...), que nun fusse capaç de falar de computadores, de telemobles ou anترنت, ou seia, para falar daquilho que ye l mais amportante de la nuossa eijistênciá, que ye l nuosso cordon que mos prende al mundo, comunicando cun las outras personas, essa lhéngua stá cundenada a zaparecer.

La lhéngua pertence al mais fondo de cada ser houmano. Daprendemos-la de maneira anstitiba, sien mos darmos de cunta, i nunca mais mudamos las sous feiçones oureginárias. Cun eilha fazemos ua biaige que bai deçde l mais fondo de la nuossa alma, l poço onde cada die la nuossa mente forma cientos de frases cun admirable criatibidade, até chegar a la comunidade onde bibimos. Falar ye cumo resfolegar. Fazemos-lo sien pensar nisso, anquanto bamos fazendo outras cousas. Mas se mos falta l aire sentimos que mos afogamos, que mos morremos – sien precisar de ningua lhagona onde seíamos atirados – i todo l restro deixa de tener amportância. La fala ye tan amportante pa la nuossa mente cumo l aire que resfolgamos. (...)

Indas que mos pareça que nun ye assi, quando miramos l ror de lhéguas faladas an ne mundo, todas eilhas ténen ua gramática i un fondo mui parecido. Ye cumo se na cabeça de todos ls ninos houbisse ua forma de barro brando, i an riba del la lhéngua fusse deixando las sues marcas. (...)

Ye nisto qu'acredítan ls lhenguístas que ban als poucos recompunendo essa pieça de barro oureginal, inda mui pouco conhecida. Mas quando stubir completa há-de mostrar muitas cousas subre l'eidentidade houmana, daquilho que muitos cháman l nuosso genoma lhenguístico.

Mas a par de l camino de ls chamados ounibersales lhenguísticos, que mos dízen que todas las lhénguas son eiguales, hai outro camino que bai a la par, cun ua lharga stória, daqueilhes que s'acupórun a tentar mostrar nó l qu'ounie las lhénguas mas aquilo que neilhas ye diferente. I deiqui até dezir qu'uas son melhores qu'outras, ou qu'esta sirbe melhor para dezir ciertas cousas, fui un passo mui fácel de dar. Assi naciú l mito de las lhénguas nacionales que, por stáren juntas al poder, atirórun cun las outras pa l buraco dialetal. Na maior parte de las bezes, al menos an Pertual i Spanha, fui un camino que lhebou tiempo i nun fui ampuosto por lei (cumo acunteciu an Fráncia qu'an 1539 decretou que la lhéngua falada na region *Île de France* era francés i todas las outras éran *patois*). Debagarico, ua lhéngua fui-se amponendo a las outras porque era la lhéngua de l poder, aqueilha que melhor permitie chegar a el ou comunicar cun mais giente. Assi s'ampeçou, mesmo sien querer, a caçoar de la maneira de falar d'alguns i a dezir qual era la maneira cierta de falar.

Durante muitos anhos mos fúrun dezindo que *falábamos atrabessado, arrebesado*, que *falábamos mal*, ou qu'*éramos palhantres*. D'ua maneira ou doutra nun falábamos bien pertués. Hai quien, inda hoije, nun tenga salido de l buraco scuro i perdura an dezir que l mirandés nun sirbe para falar d'outras cousas a nun ser l mundo rural, yá el mais zbarrulhado que la lhéngua.

EXERCÍCIO

Tomando como referência os dois últimos parágrafos do texto acima, arrole as semelhanças e diferenças entre o mirandês e o português.

Guia de leitura texto 5: O que é uma língua

- (1) Aponte o critério para se diferenciar línguas e variedades.
- (2) Por que o português do Brasil (PB) e o Português europeu (PE) são apenas variedades de uma mesma língua?
- (3) Quais as causas das dificuldades para o estabelecimento do número de línguas faladas no mundo?
- (4) Cite duas limitações de se definir uma nação por um território e uma única língua.
- (5) Quais as consequências de se entender a diversidade linguística como uma riqueza e não um problema?
- (6) O que é o ETS148 de 1992?
- (7) Explique a situação do bretão na França.
- (8) Fronteiras nacionais implicam em fronteiras linguísticas? Justifique.
- (9) Qual é a situação do Brasil, atualmente, em relação à diversidade linguística?
- (10) Qual era a situação do Brasil nos séculos XVI e XVII?
- (11) Quais as causas de o português se impor no Brasil, finalmente, no século XVIII?
- (12) Indique a importância da LDB para os povos indígenas brasileiros.
- (13) Justifique: “a diversidade linguística constitui uma parte importante do patrimônio nacional, merecendo os mesmos cuidados que o patrimônio natural e material.” (p. 69)
- (14) Resuma as definições de *língua* e *dialeto* de Mattoso Câmara (s/d).
- (15) Aponte as duas possibilidades de estudo de uma língua.
- (16) Por que falantes de línguas diferentes têm dificuldades e habilidades diferentes em relação à produção e percepção dos sons?
- (17) Por que é difícil aprender uma língua estrangeira?
- (18) Dê outros exemplos de diferenças diatópicas de nível fonético/fonológico no PB.
- (19) Indique os morfemas das palavras *bambini* (italiano), *meninos*, *amava* e *é*.
- (20) O que são alomorfos?
- (21) Como a derivação pode ser feita? Exemplifique.
- (22) De acordo com o texto, qual a definição de *a/gramatical*?
- (23) Explique: “quanto mais se reduz a marcação morfológica das pessoas verbais, tanto mais o uso dos pronomes se torna obrigatório.” (p. 78)
- (24) Defina e exemplifique gramaticalização.
- (25) O que é um diassistema?
- (26) Dê um exemplo de cada tipo de variação indicado na página 79.
- (27) Por que a diacronia não é representada na figura 1 da página 80?
- (28) Por que o círculo da diamesia está pontilhado na figura 1?
- (29) O que é repertório de um falante?
- (30) Dê exemplos de traços que nos permitem diferenciar os sistemas dentro de um diassistema.

TEXTO 6

AS LÍNGUAS DO MUNDO

Tommaso Raso e Jacyntho Lins Brandão

Das quase sete mil línguas hoje em uso no mundo, o número de falantes varia de mais de 1 bilhão, no caso do chinês, a não mais que poucas centenas ou dezenas, como nos casos do antakarinya, falado por cerca de 50 pessoas na Austrália; o omótico, com duas dezenas de falantes no Quênia; o katukina, língua falada no estado do Acre, no Brasil, por cerca de 300 pessoas; o han, do Alasca, com 40 falantes; o yevanic ou judeo-grego, em Israel, com 50; o helenorromeno ou romeno-grego, da Grécia, com 30 – e inúmeras outras.

Na tabela abaixo você encontra uma relação das línguas faladas como língua materna por mais de 4 milhões de pessoas, o que se apresenta na coluna “estimativa 1”, conforme a edição de 2009 do *Ethnologue* (cf. LEWIS, 2009, p. 20-26). Já a coluna “estimativa 2” mostra a soma do número de falantes nativos de cada uma dessas línguas ao número de falantes que as usam como segunda língua (os dados dessa segunda estimativa são bastante fluidos e, quando não há informações, o número da estimativa 1 é repetido).

A tabela se organiza na ordem decrescente da estimativa 1, devendo ser observado que, caso o critério fosse a estimativa 2, a ordenação se modificaria em vários pontos. Por exemplo, o inglês passaria do terceiro lugar para o segundo, o árabe se colocaria antes do espanhol e o russo ultrapassaria o português e o bengali. A diferença entre as duas colunas se explica em grande parte em vista de processos históricos hegemônicos de natureza cultural, religiosa ou política, que voltaremos a examinar com mais detalhes adiante.

Você encontra ainda, na segunda coluna da tabela, a família a que cada língua pertence e, quando é o caso, os grupos e subgrupos, o que lhe dará uma primeira ideia sobre quais famílias linguísticas são predominantes, considerado o número de falantes. Tendo em vista, por exemplo, as dez primeiras posições, você observará que há sete línguas indo-europeias, e apenas três pertencentes a outras famílias, um fato sobre o qual também voltaremos a refletir. Note ainda como, das línguas indígenas da América, apenas o quêchua (com 10,1 milhões de falantes nativos) e o guarani (com 4,9 milhões) integram nossa lista.

Finalmente, observe que a lista apresenta, no total, cento e quarenta e seis línguas, ou seja, menos de 2% das quase sete mil de que se tem notícia. Isso indica que mais de 80% das línguas do mundo são faladas por comunidades de menos de 4 milhões de pessoas.

TABELA 1
Línguas mais faladas

Língua	Família	Pais de base	Estimativa 1 (Língua materna)	Estimativa 2 (Língua materna + segunda língua)
Chinês	Sino-Tibetano	China	1.213.000.000	1.393.000.000
Espanhol	Indo-Europeu	Espanha	329.000.000	410.000.000
Inglês	Indo-Europeu	Reino Unido	328.000.000	508.000.000
Árabe	Afro-Asiático	Arábia Saudita	221.000.000	420.000.000
Hindi	Indo-Europeu	Índia	182.000.000	487.000.000
Bengali	Indo-Europeu	Bangladesh	181.000.000	210.000.000
Português	Indo-Europeu	Portugal	178.000.000	200.000.000
Russo	Indo-Europeu	Rússia	144.000.000	280.000.000
Japonês	Japonês	Japão	122.000.000	130.000.000
Alemão	Indo-Europeu	Alemanha	90.000.000	130.000.000
Javanês	Austronésio	Indonésia	84.600.000	89.600.000
Lahnda	Indo-europeu	Paquistão	78.300.000	78.300.000
Telugu	Dravídico	Índia	69.800.000	85.000.000
Vietnamita	Austro-asiático	Vietnam	68.600.000	85.000.000
Marathi	Indo-Europeu	Índia	68.100.000	71.000.000
Francês	Indo-Europeu	França	67.800.000	130.000.000
Coreano	Coreano	Coreia	66.300.000	70.000.000
Tamil	Dravídico	Índia	65.700.000	70.000.000
Italiano	Indo-Europeu	Itália	61.700.000	61.700.000
Urdu	Indo-Europeu	Paquistão	60.600.000	104.000.000
Turco	Altaico	Turquia	50.800.000	90.000.000
Gujarati	Indo-Europeu	Índia	46.500.000	46.500.000
Polônês	Indo-Europeu	Polônia	40.000.000	50.000.000
Malay	Austronésio	Malásia	39.100.000	39.100.000
Bhojpuri	Indo-Europeu	Índia	38.500.000	38.500.000
Awadhi	Indo-Europeu	Índia	38.300.000	38.300.000
Ucraniano	Indo-Europeu	Ucrânia	37.000.000	47.000.000
Malaiala	Dravídico	Índia	35.900.000	48.000.000
Kannada	Dravídico	Índia	35.300.000	64.000.000
Mahithili	Indo-Europeu	Índia	34.700.000	45.000.000
Sundanês	Austronésio	Indonésia	34.000.000	40.000.000
Burmese	Sino-Tibetano	Miamar	32.300.000	42.000.000
Oriya	Indo-Europeu	Índia	31.700.000	32.000.000
Persa	Indo-Europeu	Irã	31.400.000	100.000.000
Marwari	Indo-Europeu	Índia	31.100.000	31.100.000
Panjabi	Indo-Europeu	Índia	28.200.000	28.200.000
Filipino	Austronésio	Filipinas	28.000.000	50.000.000
Bhojpuri	Indo-Europeu	Índia	26.000.000	126.000.000
Hauçá	Afro-Asiático	Nigéria	25.000.000	25.000.000
Tagalog	Austronésio	Filipinas	23.900.000	85.000.000
Romeno	Indo-Europeu	Romênia	23.400.000	30.000.000
Indonésio	Austronésio	Indonésia	23.200.000	140.000.000

Faculdade de Letras da UFMG - Fundamentos de Linguística Comparada - Apostila

Holandês	Indo-Europeu	Holanda	21.700.000	25.000.000
Sindhi	Indo-Europeu	Paquistão	21.400.000	28.000.000
Thai	Tai-Kadai	Tailândia	20.400.000	28.000.000
Pashto	Indo-Europeu	Paquistão	20.300.000	25.000.000
Uzbek	Altaico	Uzbesquistão	20.300.000	25.000.000
Rajasthani	Indo-Europeu	Índia	20.000.000	25.000.000
Arzerbaijano	Altaico	Irã	19.400.000	39.000.000
Iorubá	Nigero-Congolês	Nigéria	19.400.000	21.000.000
Igbo	Nigero-Congolês	Nigéria	18.000.000	18.000.000
Amhárico	Afro-Asiático	Etiópia	17.500.000	34.000.000
Chhattisgarhi	Indo-Europeu	Índia	17.500.000	17.500.000
Oromo	Afro-Asiático	Etiópia	17.300.000	26.000.000
Assamês	Indo-Europeu	Índia	16.800.000	16.800.000
Servo-Croata	Indo-Europeu	Sérvia	16.400.000	17.000.000
Curdo	Indo-Europeu	Iraque	16.000.000	30.000.000
Cebuano	Austronésio	Filipinas	15.800.000	30.000.000
Sinhalês	Indo-Europeu	Sri Lanka	15.600.000	15.600.000
Rangpuri	Indo-Europeu	Bangladesh	15.000.000	15.000.000
Thai Norte-oriental	Tai-Kadai	Tailândia	15.000.000	15.000.000
Zhuang	Tai-Kadai	China	14.900.000	15.000.000
Malgaxe	Austronésio	Madagascar	14.700.000	14.700.000
Nepali	Indo-Europeu	Nepaço	13.900.000	30.000.000
Somali	Afro-Asiático	Somália	13.900.000	13.900.000
Khmer	Austro-Asiático	Camboja	13.600.000	15.000.000
Madurês	Austronésio	Indonésia	13.600.000	14.000.000
Bávaro	Indo-europeu	Austria	13.300.000	13.300.000
Grego	Indo-Europeu	Grécia	13.100.000	30.000.000
Chittagoniano	Indo-Europeu	Bangladesh	13.000.000	13.000.000
Haryanvi	Indo-Europeu	Índia	13.000.000	13.000.000
Magahi	Indo-Europeu	Índia	13.000.000	13.000.000
Decan	Indo-Europeu	Índia	12.800.000	12.800.000
Húngaro	Urálico	Hungria	12.500.000	14.000.000
Fula	Nigero-Congolês	Senegal	12.300.000	12.300.000
Catalão	Indo-Europeu	Espanha	11.500.000	12.000.000
Shona	Nigero-Congolês	Zimbabwe	10.800.000	17.000.000
Sylheti	Indo-Europeu	Bangladesh	10.300.000	10.300.000
Zulu	Nigero-Congolês	África do Sul	10.300.000	25.000.000
Quêchua	Ameríndio	Peru	10.100.000	15.000.000
Kanauji	Nigero-Congolês	Índia	9.500.000	9.500.000
Tcheco	Indo-Europeu	Rep. Tcheca	9.500.000	12.000.000
Lombardo	Indo-Europeu	Itália	9.100.000	9.100.000
Búlgaro	Indo-Europeu	Bulgária	9.000.000	9.000.000
Uyghur	Altaico	China	8.800.000	8.800.000
Nyanja	Nigero-Congolês	Malawi	8.700.000	8.700.000
Bielorusso	Indo-Europeu	Bielorússia	8.600.000	10.000.000
Akan	Nigero-Congolês	Gana	8.300.000	10.000.000

Fonte: Lewis, 2009, p. 20-26.

1 Áreas e famílias linguísticas

Como apresentados acima, os dados dizem pouco em termos geográficos e históricos – no sentido de que possamos compreender por que as línguas são tantas e são faladas em proporção tão desigual. Para entender um pouco mais essa realidade, as agruparemos a partir de dois critérios: sua distribuição por áreas; sua distribuição por famílias.

A tabela 2 organiza os dados quantitativos relativos a cinco grandes áreas linguísticas: África, Américas, Ásia, Europa e Pacífico. Ressalte-se que estamos falando de áreas linguísticas e não geográficas, ainda que os espaços possam coincidir. O que nos interessa, neste caso, é o local de procedência das línguas, mesmo que elas sejam faladas em espaços geográficos distintos. Os milhões de falantes nativos de português, espanhol, inglês e francês nos países americanos, por exemplo, são computados na área linguística da Europa, não na das Américas, uma vez que essas línguas tiveram origem naquele continente. O mesmo vale para os falantes nativos de inglês da Austrália, computados na área linguística da Europa e não do Pacífico – e assim por diante.

Tendo em vista esses parâmetros, os números são os seguintes:

TABELA 2
Línguas e falantes das cinco grandes áreas linguísticas mundiais

Área de origem	Línguas		Falantes			
	Número de línguas	Percentual	Número de falantes	Percentual	Média	Ponto médio
África	2.110	30,5%	726.453.403	12,2%	344.291	25.200
América	993	14,4%	50.496.321	0,8%	50.852	2.300
Ásia	2.322	33,6%	3.622.771.264	60,8%	1.560.194	11.100
Europa	234	3,4%	1.553.360.941	26,1%	6.638.295	201.500
Pacífico	1.250	18,1	6.429.788	0,1%	5.144	980
Totais	6.909	100%	5.959.511.717	100%	862.572	7.560

Fonte: Lewis, 2009, p. 19.

Alguns dados são significativos:

(a) Como você pode observar na segunda coluna, a área com menos línguas é a Europa, com 234, enquanto a Ásia e a África são as duas regiões com mais línguas, respectivamente 2.322 e 2.110.

(b) Assim, as línguas europeias representam apenas 3,4% das línguas do mundo,

enquanto tanto as asiáticas quanto as africanas apresentam percentuais quase dez vezes maiores.

(c) Considerando-se o número de falantes, a situação se modifica bastante: as línguas europeias, que constituem apenas 3,4% das línguas do mundo, são faladas por mais de 26% da humanidade; ao contrário, o Pacífico, que concentra mais de 18% do total de línguas, responde por apenas 0,1% do de falantes.

(d) Note-se que as línguas europeias e asiáticas juntas representam 37% das línguas do mundo, mas são faladas por quase 87% da população mundial. Já as línguas das Américas e do Pacífico juntas, apesar de representarem 32,5% do número de línguas, não passam de 2% quanto se trata do número de falantes.

(e) À primeira vista, a Ásia apresenta uma situação bastante equilibrada, com os maiores percentuais em termos de línguas e de falantes, o mesmo podendo se dizer com relação à África, com o segundo maior número de línguas e o terceiro maior número de falantes.

(f) Contudo, quando se considera, na penúltima coluna da tabela, a média de falantes por língua, observa-se que a média europeia, de mais de 6,3 milhões de falantes por língua, é mais de quatro vezes a das línguas asiáticas e quase vinte vezes a das africanas – bem como mais de cento e trinta vezes a das línguas americanas e mais de mil e quinhentas vezes a das línguas do Pacífico!

Evidentemente que números, médias e percentuais representam não mais que dados estatísticos de ordem geral, significativos em si, mas que não contemplam toda a diversidade de situações. Se na área linguística europeia a média de falantes por língua é de mais de 6 milhões, como vimos, há nela línguas faladas por muito menos pessoas, pois esse índice tem como referência, num extremo, o inglês, com seus 340 milhões de falantes nativos, e, no outro, o Helenorromeno, com seus 40.

Por isso é importante considerar o índice da última coluna, o “ponto médio” relativo a cada área linguística, assim estabelecido: o rol de línguas faladas em cada área é dividido em duas metades, observando-se, na linha de corte, o número de falantes da língua que ocupa essa posição. Assim, das 234 línguas europeias, 117 contam com mais de 220 mil falantes e 117 com menos que isso; na África, 1.055 línguas têm mais de 25 mil falantes e o restante menos; na Ásia, 1.161 línguas são faladas por mais de 11 mil pessoas e outras 1.161 por menos; nas Américas, 496 línguas têm mais de 2.300 falantes e outro tanto menos que isso; finalmente, no Pacífico, 625 línguas têm mais de 980 falantes e outras 655 menos que 980.

Como se vê, as línguas do Pacífico estão em claro risco de extinção, considerando-se tanto a média de falantes por língua, que é de 4.675 pessoas, quanto o ponto médio, com 800 falantes. Com efeito, costuma-se considerar que, em longo prazo, uma língua falada por

menos de 100 mil pessoas está condenada ao desaparecimento. Uma projeção de ordem geral indica que 55% das línguas do mundo contam com menos de 10 mil falantes, um quarto delas com menos de mil. É provável, portanto, que algumas das línguas do Pacífico acima computadas já tenham desaparecido desde a publicação desses dados. Mas também nas Américas a situação é precária: a média é de pouco mais de 50 mil falantes por língua (e recorde-se que apenas o quêchua e o guarani somam 15 milhões, ou seja, quase 30% do total de 50 milhões de pessoas que falam línguas americanas), o ponto médio sendo de pouco mais de dois mil falantes.

Outro modo de analisar a situação linguística do mundo é agrupar as línguas por famílias, o que pode nos fornecer uma perspectiva histórica complementar à relativa às áreas linguísticas. O que se modifica, neste caso, é que, por exemplo, algumas línguas da área europeia, como o húngaro e o finlandês, da família urálica, não serão agrupadas com o grupo principal nesse espaço, o indo-europeu, ao passo que uma parte das línguas da Ásia, como o persa, o hindi e o bengali, o serão.

Os dados relativos às seis maiores famílias são os que se mostram na tabela a seguir:

TABELA 3
Famílias linguísticas

Família	Línguas		Falantes			
	Número de línguas	Percentual	Número de falantes	Percentual	Média	Ponto médio
Afro-asiática	353	5,11%	339.478.607	5,93%	961.696	20.151
Austronésia	1.246	18,03%	311.740.132	5,45%	250.193	3.384
Indo-europeia	430	6,22%	2.562.896.428	44,78%	5.960.224	150.000
Nigero-cordofoniana	1.495	21,63%	358.091.103	6,26%	239.526	26.000
Sino-tibetana	399	5,77%	1.275.531.921	22,28%	3.196.822	18.686
Trans-Nova-Guiné	561	8,12%	3.359.894	0,06%	5.989	1.245
Outras famílias	2.428	35,13%	872.763.125	15,25%	359.457	-

Observe que a família indo-europeia, com apenas 6% do total de línguas (430) atinge quase 45% da população mundial, enquanto a família trans-Nova-Guiné, que apresenta um número de línguas bem próximo, embora ligeiramente maior (561 línguas, ou seja, 8,12% do total), abrange só 0,06% dos falantes. Juntas, a família indo-europeia e a sino-tibetana, que somam 829 línguas (11,99% das línguas do mundo), cobrem quase 70% da população.

O que os dados das duas tabelas têm em comum é a concentração de grandes línguas no espaço da Europa e da Ásia. Apontam também em comum a tendência à diminuição do número de falantes das línguas dos outros continentes e aumento da quantidade de pessoas que falam línguas euroasiáticas – o que a média de falantes por língua e o ponto médio

expressam com bastante clareza –, a exceção ficando por conta do grupo banto, da família nigero-cordofoniana, pelas razões que veremos mais à frente.

2 Fatores históricos, culturais e políticos

Como se explicam esses fatos? Antes de tudo, devemos considerar que o destino das línguas se liga ao destino dos povos que as falam. Hoje, por exemplo, o inglês é tão difundido porque os Estados Unidos são muito poderosos. Mas por que os indo-europeus e os sino-tibetanos se tornaram tão poderosos, enquanto os aborígenes australianos e os ameríndios brasileiros não? E mais: por que foram os povos e as línguas europeus que conquistaram o mundo e não os ameríndios que conquistaram a Europa e a Ásia?

Uma das explicações possíveis e que parece bastante razoável está no fato de que aonde a agricultura chegou a tempos mais remotos e teve a possibilidade de espalhar-se, inevitavelmente as comunidades agrícolas acabaram se estendendo às custas das de caçadores e coletores. Estas últimas foram assimiladas pelas primeiras ou dizimadas – ou, ainda, o que ocorreu em alguns casos, ficaram confinadas em pequenas áreas marginais (cf. DIAMONDS, 2006, p. 83-192; CAVALLI-SFORZA, 2003, p. 127-176). O mesmo destino acompanhou as línguas dessas comunidades, que acabaram desaparecendo – ou, quando confinadas, reduzindo-se a poucos falantes, como parece ter acontecido com o basco, uma língua europeia isolada, provavelmente falada no espaço em que ainda se encontra anteriormente à ocupação da Europa por falantes de línguas indo-europeias. Evidentemente é muito difícil determinar como se desenrolou a história linguística do mundo, pois isso supõe embrenhar por eras muito antigas.

Na esteira de Diamonds e Cavalli-Sforza, vamos dar um passo atrás para verificar a importância da agricultura nos destinos da humanidade, acompanhando seu surgimento. A espécie humana, mas ainda não com as características cognitivas que tem, surgiu provavelmente na África Oriental, por volta de sete milhões de anos atrás. Por volta de um milhão de anos, essa espécie, chamada *homo erectus*, depois de ter-se expandido pela África, ocupou o mundo todo, menos a Austrália e as Américas. Mas foi o *homo sapiens*, primeiro o neanderthaliano e depois a espécie que podemos chamar de nossa, que ocupou o mundo todo. A nossa espécie se desenvolveu cerca de 50 mil anos atrás, mas os neanderthalianos, que foram completamente eliminados, já eram bons caçadores.

Não se pode ter certeza da cronologia da ocupação do mundo em épocas tão remotas, mas uma boa hipótese é que o homem saiu da África há cerca de um milhão de anos para ocupar o Oriente Médio, depois se dividiu em dois grupos: um foi para o leste e ocupou toda a Ásia do Sul; outro foi para noroeste e ocupou a Europa, por volta de 500 mil anos atrás.

Somente a nossa espécie, surgida faz aproximadamente 50 mil anos, conseguiu ocupar a Ásia do Norte, cerca de 20.000 a.C., a Austrália, cerca de 40.000 a.C., e as Américas. Pelo que se sabe, ainda que os dados não sejam tão seguros, o homem moderno chegou à América do Norte apenas por volta de 12.000 a.C. e ocupou rapidamente todo o continente, atingindo o extremo sul em 10.000 a.C. A ocupação das Américas foi, portanto, muito tardia, mas, ao mesmo tempo, extremamente rápida.

Esse quadro permite responder a outra pergunta importante: por que o continente sul-americano, que possui condições ambientais tão parecidas às da África, não possui grandes mamíferos. Parece sensata a seguinte hipótese: o homem, na África, desenvolveu suas próprias características cognitivas e suas próprias capacidades de caçador apenas aos poucos; e, enquanto o homem desenvolvia essas capacidades, os animais aprendiam a defender-se. Ao contrário, quando o homem chegou à América do Sul, depois de sete milhões de anos, já era o homem da nossa espécie, cognitivamente muito mais evoluído e treinado por milênios de caça. Isso fez com que ele exterminasse rapidamente todos os grandes mamíferos, os quais estavam despreparados para lidar com um predador tão poderoso. Foi o primeiro desastre ecológico da humanidade. Isso terá consequências importantes para a agricultura.

Sabemos que a agricultura surgiu no Oriente Médio cerca de 10.000 a.C. (mais ou menos quando o homem chegava nas Américas). Portanto, o homem, que surgiu cerca de sete milhões de anos atrás, passou apenas 0,5% da própria história com a agricultura – e 99,5% sem agricultura. Ou seja, se consideramos a história da humanidade como um dia, podemos dizer que a agricultura surgiu por volta de dez ou quinze minutos antes da meia-noite. Seria estranho pensar somente nos últimos quinze minutos e não dar nenhuma atenção às 23 horas e 45 minutos que vieram antes. Mesmo se nos interessássemos somente pelos humanos da nossa espécie, aquela de todo cognitivamente desenvolvida e que acabou com os neanderthalianos, partindo da agricultura perderíamos 80% de sua história. É preciso considerar outro aspecto: se a agricultura se desenvolveu no Oriente Médio por volta de 10.000 a.C., nas Américas isso aconteceu somente em 1.500 d.C., na Austrália, como em outros muitos locais, bem depois – e em alguns poucos lugares, como em algumas tribos indígenas da Amazônia, isso ainda não se verificou. Portanto, não entender as condições de vida do homem caçador significaria não entender os índios brasileiros até a chegada dos portugueses e as condições de algumas tribos até hoje.

Os caçadores-coletores viviam de maneira muito diferente dos agricultores, ou seja, de nós. Eles se alimentavam, obviamente, com o que a natureza oferecia: os frutos naturalmente produzidos pela vegetação nativa e os animais que eles conseguiam caçar. Periodicamente, mudavam de lugar, sendo nômades. Isso era necessário para fazer com que uma área, cujos produtos já tivessem sido desfrutados, pudesse regenerar-se. Eles não podiam guardar nada,

porque os frutos e as carnes, depois de poucos dias, apodrecem. De igual modo, não podiam formar grupos muito numerosos, porque a natureza não forneceria alimento para muitas pessoas e seria difícil transportar as crianças quando se mudava de área. Assim, as comunidades de caçadores-coletores eram formadas por um número pequeno de indivíduos e ocupavam uma área muito grande. Existia um equilíbrio natural entre o espaço geográfico-alimentar e o número de indivíduos que nele podia ser hospedado.

Obviamente, a vida de uma comunidade de caçadores-coletores era muito simples: a atividade principal era destinada à busca de comida, com as mulheres se dedicando à coleta e os homens à caça. Eles tinham alguns rituais, fabricavam alguns objetos de pedra e madeira, não possuindo uma verdadeira hierarquia interna, pois todos desempenhavam as mesmas poucas atividades. De vez em quando, entravam em contato com comunidades vizinhas, que moravam, de toda maneira, bem longe, por poucas razões: para trocar parceiras em casamento; para trocar produtos, ou seja, para um comércio extremamente incipiente; para fazer a guerra pelo domínio do território e de seus recursos alimentares. De vez em quando, se a comunidade crescia demais, uma parte se separava para buscar recursos. Em suma, a população não tinha muitas possibilidades de crescer. O limite de recursos controlava a população e, quando isso não era suficiente, a guerra com os vizinhos fazia o resto.

Quando surge a agricultura, esse equilíbrio se quebra definitivamente. O homem pode plantar seus produtos, acumular bens, morar sempre no mesmo lugar. Manipulando a natureza, ele consegue uma quantidade de produtos superior ao necessário para a sobrevivência imediata e não precisa mais limitar o número de membros da comunidade. Deixando de ser nômade, não tem mais problemas de transporte e pode ter mais filhos e riquezas. O crescimento traz, assim, várias consequências, esquematizadas no quadro que segue:

QUADRO 1
Organizações populacionais humanas

Tipo de organização	Bando	Tribo	Tribo centralizada	Estado
Dimensão				
Número de pessoas	Dezenas	Centenas	Milhares	Mais de 50.000
Assentamentos	Nômade	Vilarejo	Vilarejos	Cidades
Relações	Parentesco	Clã	Classe social	Classe social
Etnias e línguas	Uma	Uma	Uma	Uma ou mais
Formas de governo				
Decisões	Igualitárias	Influenciadas por um chefe	Centralizadas	Centralizadas
Burocracia	Não existe	Não existe	Simple	Complexa
Controle do poder	Não existe	Não existe	Sim	Sim
Solução de conflitos	Informal	Informal	Centralizada	Leis, juízes
Hierarquia dos assentamentos	Não existe	Não existe	Vilarejo central	Cidade central
Religião				
Justifica a cleptocracia?	Não	Não	Sim	Sim > não
Economia				
Agricultura e criação animal?	Não	Não > sim	Sim > intensiva	Intensiva
Divisão trabalho?	Não	Não	Não > sim	Sim
Trocas?	Recíprocas	Recíprocas	Redistribuídas (impostos)	Redistribuídas (impostos)
Organização social				
Estratificação?	Não	Não	Sim (clã)	Sim
Escravidão?	Não	Não	Em pequena escala	Em grande escala
Produção de bens de luxo para as elites?	Não	Não	Sim	Sim
Edifícios públicos?	Não	Não	Não > sim	Sim
Escrita?	Não	Não	Não	Frequentemente

Fonte: Diamonds, 2006, p. 268.

O que se pretende é não mais que esquematizar os possíveis estágios em que se dá a passagem, graças à agricultura, das comunidades em bandos até as sociedades modernas. Observe alguns aspectos:

1. Passa-se de uma organização igualitária para uma estratificação em classes sociais. O

fato é que, com o crescimento populacional, se torna mais conveniente diferenciar e especializar o trabalho, ou, noutros termos, quando a comunidade começa a tornar-se complexa, parece melhor que cada qual conheça melhor um tipo de trabalho, gerando implementações. Mas, naturalmente, nem todos os trabalhos estão no mesmo nível: se todo mundo pode ser agricultor, não todos podem ser bons guerreiros e, menos ainda, bons políticos. Assim, gera-se uma estratificação social baseada, em princípio, nas capacidades e, em seguida, no interesse. Essa divisão se auto-alimenta, pois quem está numa situação melhor tem acesso a bens e oportunidades que servem para perpetuar a sua superioridade e tornar maior a sua própria prosperidade.

2. Somente as sociedades complexas conhecem a escrita, a qual surge no Oriente Médio, cerca de 3.500 a.C., pressupondo a agricultura. A escrita potencializa enormemente as atividades de uma sociedade e contribui para uma mudança de foco e de cognição. Ela permite superar os limites de tempo e espaço na comunicação, liberar a memória e fixar as normas, as tradições e a história em uma versão única, com a qual todos podem se confrontar. Ao contrário, as culturas exclusivamente orais permitem que a história seja transformada por quem é o depositário dela.

3. As indicações sobre as línguas, no quadro, devem ser interpretadas. Quando se diz que, nos primeiros estágios, se possui uma língua, e, no último, várias, entende-se que as sociedades complexas hospedam imigrantes que falam outras línguas – bem como que, nelas, se promove o aprendizado de línguas diferentes.

Por que este longo discurso sobre os efeitos da agricultura num curso de linguística? Porque é a agricultura que gera o quadro mostrado nas tabelas anteriores. Os povos de agricultura mais antiga cresceram rapidamente e foram dominando territórios cada vez mais amplos, absorvendo ou exterminando os caçadores-coletores que encontravam no caminho. Seu predomínio se fez destruindo o meio-ambiente de onde os caçadores-coletores tiravam seu alimento, já que a agricultura destrói a natureza nativa, para moldá-la a seus objetivos; seu predomínio se deu pela força dos números, já que as comunidades de caçadores-coletores eram sempre pequenas e as organizações agrícolas podiam crescer sem limites; seu predomínio se deu pela superioridade tecnológica, já que, com o acúmulo de bens, sua organização social complexa e a especialização do trabalho, as comunidades agrícolas criaram exércitos profissionais, descobriram e desenvolveram armas superiores, treinaram os cavalos para a guerra e criaram máquinas capazes de multiplicar a força destruidora das armas primitivas.

No momento em que as organizações agrícolas absorvem ou destroem os bandos de caçadores, eliminam automaticamente suas línguas. Num contexto de muitos bandos de caçadores-coletores em equilíbrio entre si, temos muitas línguas também em equilíbrio.

Quando, ao contrário, uma sociedade agrícola cresce, é sempre a mesma língua que cresce. Portanto, um contexto de comunidades de caçadores-coletores gera um contexto de muitas línguas, cada uma delas com poucos falantes. Um contexto agrícola faz com que a uma organização social corresponda uma única língua com muitos falantes. Os locais de agricultura mais antiga, Oriente Médio, Europa e Ásia, são aqueles onde se encontram também as mais antigas organizações estatais (Egito, Mesopotâmia, China, Pérsia, Grécia, Roma etc) e as famílias linguísticas predominantes (principalmente a sino-tibetana, a indo-europeia e a afro-asiática). Os locais onde a agricultura chegou mais tarde ou ainda não chegou, são onde se encontram línguas faladas por pequenos grupos.

Como vimos, a Europa é, de longe, o continente com a menor diferenciação linguística. Mas, ao mesmo tempo, é famosa exatamente pela impressão que dá ao viajante de que com frequência há mudança de línguas. Quem se desloca por aquele continente percebe isso claramente, enquanto quem viaja pelo Brasil ou pelos Estados Unidos tem a sensação que esses países, tão vastos quanto a Europa, são monolíngues, sem suspeitar que a diferenciação linguística americana e brasileira é muito maior do que a europeia. Como explicar essa contradição? Com a colonização iniciada no século XVI, os europeus empurraram as comunidades locais que sobreviveram aos processos de extermínio e assimilação para lugares impenetráveis, como a floresta amazônica, ou para pequenas reservas. Portanto, na maior parte dos países da América há uma língua que abrange a quase totalidade do território e dos falantes e muitas línguas faladas em pequenos enclaves pouco acessíveis e com pouquíssimos falantes.

Como aconteceu esse processo? Como foi possível que poucos milhares de espanhóis, ingleses e portugueses destruíssem, em poucas décadas, de 90 a 95% da população ameríndia? Não foi somente nem principalmente com a força das armas, mas com uma força que nem os próprios europeus imaginavam ter: a das doenças. De fato, um bom pacote agrícola, como o euro-asiático, juntava a produção estritamente agrícola com a criação de animais domésticos, os quais, dependendo do lugar, podiam ser vários, mas quase todos parecem ter uma origem medio-oriental: a galinha, o boi, a cabra, a ovelha etc. Eles contribuía de várias maneiras: com as peles e a lã, para fazer tecidos; com a carne e os ovos, para fornecer as proteínas antes encontradas na caça; com o leite e seus derivados, para enriquecer a alimentação; com sua força de tração, para guiar o arado e melhorar as técnicas agrícolas, bem como para facilitar o transporte. Trata-se de um conjunto poderosíssimo de vantagens. Ora, como homens e animais passam a viver juntos, isso fez com que passassem a compartilhar também as doenças. Como nos países de agricultura mais antiga o contato entre homem e animais foi gradual, os agricultores desenvolveram também anticorpos contra as enfermidades transmitidas pelos animais domésticos. Quando sociedades que vinham de

milênios de agricultura, como as dos portugueses, dos espanhóis e dos ingleses, entraram em contato com as comunidades indígenas das Américas, estas não tinham nenhuma defesa contra as epidemias comuns na Europa, seus integrantes vindo a morrer de gripe, varíola e outras enfermidades. (Cf. DIAMOND, 2006, p. xx)

Precisamos agora formular e responder outra pergunta. Por que a agricultura nasceu no Oriente Médio e por que se desenvolveu tão bem na Eurásia? E ainda: por que a agricultura não nasceu ou não se desenvolveu em outros continentes com a mesma pujança que na Eurásia? Um primeiro problema é dado pelo fato de que parece que as condições de vida dos agricultores, por muito tempo, foram piores que as dos caçadores-coletores. Os dados arqueológicos mostram uma redução da estatura dos primeiros agricultores e uma diminuição na duração de suas vidas. Por que então se passou da caça à agricultura? Como normalmente acontece na história, parece que o motivo da mudança foi uma crise. Neste caso, uma crise relativa às condições ambientais, como uma redução forte e imprevista dos animais e dos frutos por motivos desconhecidos. A população teve então de inventar um novo sistema de sobrevivência e começou a cultivar, com técnicas rudimentares, frutos que certamente eram de qualidade e tamanho muito diferentes daqueles que temos agora, após treze mil anos de domesticação. Mas aos poucos o novo sistema foi-se aperfeiçoando e revelou potencialidades impensáveis, como vimos pelas consequências que trouxe. Portanto, nos lugares onde não ocorreram as mesmas dificuldades, as populações não tiveram motivos para mudar os hábitos tradicionais.

Observe-se que a agricultura surgiu, mesmo que não em épocas tão antigas, também em outros locais: na Amazônia, no vale do Mississipi, no nordeste dos Estados Unidos, nos Andes, nas montanhas de Papua-Nova Guiné, na África Oriental, na região entre Camarões e a Nigéria. O que fez com que esses lugares não produzissem civilizações poderosas, como as da China, da Europa e do Oriente Médio? É provável, como crê Diamond, que sobretudo devido também a questões ambientais. De fato, o pacote desenvolvido na China e no Oriente Médio tinha dois motivos para prevalecer fortemente.

O primeiro, que nas Américas e em Papua-Nova Guiné praticamente não existiam mais grandes mamíferos, todos massacrados, provavelmente, na primeira chegada do homem. A África subsaariana possuía muitos grandes mamíferos, mas nenhum domesticável: até hoje ninguém conseguiu domesticar leões, veados, girafas ou zebras; quanto aos elefantes, podem ser adestrados, mas não podem se reproduzir nesse estado. O mesmo pode ser dito, ainda que com menos evidências, com relação às plantas. O inhame da África, a batata doce de Papua-Nova Guiné e a mandioca da Amazônia não podiam competir com o trigo mediterrâneo e com o arroz chinês.

O segundo motivo parece estar na estrutura geográfica dos continentes: a Eurásia se

estende ao longo de um eixo leste-oeste e, portanto, apresenta o mesmo clima por toda a sua extensão. Além disso, não existem grandes barreiras entre o Oriente Médio, a Europa e a Ásia. Isso fez com que os pacotes agrícolas fossem transportados e se integrassem com facilidade em várias regiões. Ao contrário, a África e as Américas se estendem ao longo do eixo norte-sul, com abruptas mudanças de clima, o que torna difícil transportar de uma região para a outra a mesma planta ou mesmo animal, em vista das condições ambientais. A isso aliam-se barreiras naturais. A agricultura que surgiu nas montanhas andinas não desceu para a Amazônia, apesar da pouca distância. Na África, o deserto do Saara funcionou como uma grande barreira entre o resto do continente e o Oriente Médio, bem como a Nigéria estava separada, pela floresta equatorial, da África oriental.

Ora, nas tabelas anteriores você pôde verificar que duas famílias linguísticas, a indo-europeia e a sino-tibetana, ou seja, as maiores da Eurásia, com apenas 829 línguas (muitas das quais faladas por comunidades muito pequenas), abrangem 3,8 bilhões de falantes, uma alta percentagem da população mundial. Todavia, enquanto a família sino-tibetana está localizada numa única região da Ásia, a indo-europeia está presente em todos os continentes. Esses dados apresentam dois aspectos: (a) que a agricultura, surgida e desenvolvida nos dois continentes da Eurásia, determinou em grande parte que as grandes famílias euroasiáticas crescessem como nenhuma outra do planeta; (b) mas também algo, num certo ponto da história, determinou que somente os europeus se expandissem por outros continentes, enquanto os asiáticos permaneceram numa região definida. Isso é verdade até certo ponto, pois a também a família indonésia chegou até Madagascar, expandindo-se por uma enorme área do Oceano Índico e do Pacífico.

O que determinou que os europeus viajassem até as Américas e os chineses, com uma cultura, uma economia, um potencial militar e uma frota naval superiores, ficassem parados? Parece que a resposta está na crise que a Europa viveu com a grande peste do século XIV, a qual destruiu as populações numa fase de grande expansão econômica. O continente já tinha passado antes por uma crise gigantesca, com a queda do Império Romano, como se pode constatar por seu decréscimo em termos demográficos: se na parte da Eurásia pertencente ao Império Romano – de Portugal até o Oriente Médio, incluindo o norte da África – a população era, por volta do nascimento de Cristo, de mais de 200 milhões, com a crise que se seguiu à dissolução decresceu consideravelmente, voltando ao mesmo patamar somente mais de mil anos depois. A grande peste de metade do século XIV também produziu abrupta diminuição populacional, somente por volta de 1500 voltando a verificar-se números compatíveis com os anteriores à epidemia. Por que essas crises poderiam apontar alguma explicação para o fato de terem sido os europeus (e não os chineses) a chegar nos outros continentes? O que parece que aconteceu foi que, com a retomada do crescimento

demográfico depois do ano 1000, enquanto o continente se desenvolvia economicamente e produzia muitos bens, de repente, em 1350, grande parte da população que podia adquirir e consumir tanta produção desapareceu. Isso deu início às grandes navegações, em busca de outros mercados, principalmente na África e na Ásia. Foi assim, buscando um caminho mais curto para a Ásia, que se descobriu, por acaso, a América, tendo início a fase da colonização. A China não passou por crises semelhantes e, portanto, nunca precisou buscar outros mercados.

Voltando à tabela 1, você poderá observar como parece que todos esses fatores sugerem, pelo menos em parte, razões para o predomínio de certas famílias e áreas linguísticas – todas da Eurásia. Se tomarmos apenas as línguas hoje faladas para mais de 100 milhões de pessoas, como primeira ou segunda língua, fica claro sua procedência europeia e asiática, como se recorda na tabela abaixo:

TABELA 4
Línguas mais faladas, famílias e áreas linguísticas

Língua	Família	Área linguística	Número de falantes
Chinês	Sino-Tibetano, Chinês	Asiática	1.393.000.000
Inglês	Indo-Europeu, Germânico	Europeia	508.000.000
Hindi	Indo-Europeu, Indo-Iraniano, Índico	Asiática	487.000.000
Árabe	Afro-Asiático, Semítico	Asiática	420.000.000
Espanhol	Indo-Europeu, Itálico, Românico	Europeia	410.000.000
Russo	Indo-Europeu, Esloveno	Europeia	280.000.000
Bengali	Indo-Europeu, Indo-Iraniano, Índico	Asiática	210.000.000
Português	Indo-Europeu, Itálico, Românico	Europeia	200.000.000
Japonês	Japonês	Asiática	130.000.000
Alemão	Indo-Europeu, Germânico	Europeia	130.000.000
Francês	Indo-Europeu, Itálico, Românico	Europeia	130.000.000

Instrutivo também será considerar a distribuição geográfica das línguas atualmente ameaçadas de extinção. Considerando, dentre estas, apenas aquelas faladas por pessoas idosas e não mais transmitidas às crianças – ou seja, em estado de extinção iminente – temos a seguinte tabela:

TABELA 5
Línguas ameaçadas de extinção

Área de origem	Número total de línguas	Línguas ameaçadas de extinção	Percentual de línguas ameaçadas na área
América	1.002	170	16,96%
Pacífico	1.310	210	16,03%
Europa	239	12	5,02%
Ásia	2.269	78	3,43%
África	2.092	46	2,19%

Observe como as situações mais graves se encontram, pela ordem, na América e no Pacífico. Tomando como referência as outras áreas, podemos admitir que, considerando-se que as línguas tanto surgem quanto desaparecem, percentuais em torno de 3,5 seriam de se esperar. Na América e no Pacífico o que temos, contudo, são percentuais mais de quatro vezes maiores, o que constitui uma situação anômala, uma razão provável estando no fato de que se trata de áreas de expansão recente de línguas hegemônicas de origem europeia – no caso das Américas –, ou europeias e asiáticas – com relação ao Pacífico.

Há todavia outro aspecto a ser lembrado: nas áreas com menos línguas ameaçadas a grande extinção linguística já teve lugar há muito mais tempo – o que vale se considerarmos a expansão tanto indo-europeia, quanto sino-tibetana, que vêm se desenrolando desde eras pré-históricas. O mesmo vale para a África, em que se observa a predomínio de dois grupos linguísticos, o nigero-cordofaniano e o nilo-saariano, cujas populações dominaram vastas áreas do continente. A diferença com o que ocorre hoje com as línguas das Américas e do Pacífico é que talvez nunca o processo se tenha acelerado tanto.

Em alguns poucos casos é possível saber com precisão a data de extinção de uma língua. Um dos exemplos mais famosos é relativo ao dálmata, língua românica, procedente do latim, como o português, falada desde fins da Antiguidade no litoral e nas ilhas de onde hoje se encontram a Croácia e Montenegro. De dois de seus dialetos se possuem registros: o ragusano (cuja denominação provém de Ragusa, nome antigo da atual cidade de Dubrovnik, na Croácia), conhecido por textos datados de entre os séculos XIV e XVI, quando deixou definitivamente de ser falado, suplantado pelas línguas eslavas da região; e o velhoto, falado na ilha de Veglia (nome italiano da antiga Viklasun, atualmente chamada de Krk) até o século XIX. Com relação a este último, a data de seu desaparecimento é precisa: 10 de junho de 1898, quando morreu, vitimado por uma mina terrestre, seu último falante, Tuone Udaina Burbur (também conhecido por seu nome italiano: Antonio Udina). Consta que este Tuone Udaina, submetido a um interrogatório, se recusou a responder em italiano, falando intencionalmente em dálmata, a fim de confundir os policiais. Tendo sido chamada a atenção para esse fato, o

acadêmico italiano Matteo Giulio Bartoli, ele próprio natural de uma região próxima, a Ístria, visitou, em 1897, Tuone Udaina, registrando cerca de 2.800 palavras do velhoto, além contos e relatos da vida de seu informante. A partir, sobretudo, desse material, Bartoli escreveu um livro tratando do vocabulário, da fonologia e da gramática do dálmata, obra publicada em tradução para o alemão (Bartoli, M. G. *Das Dalmatische*. Viena: Kaiserliche Akademie der Wissenschaft, 1906). É assim que o registro dessa língua românica oriental não se perdeu, permitindo o conhecimento de fenômenos fonológicos, morfológicos e sintáticos que ela compartilha com o romeno e o vêneto, bem como de características próprias apenas a ela. (cf. VIDOS, 1973, p. 315-317).

Outro exemplo recentíssimo é o do bo, língua falada nas ilhas Andaman, no Golfo de Bengala, Índia. Conforme noticiou a Profa. Anvita Abbi, do Centro de Linguística da Universidade Jawaharlal Nehru (Nova Delhi), com a morte de Boa Sr., em Port Blair, em 26 de janeiro de 2010, aos 85 anos, desapareceu a última falante desta língua. Esta já não tinha mais com quem conversar no seu idioma materno, razão por que aprendera também outras línguas andamanesas da mesma ilha, além do hindi, mas era o único membro de sua etnia que se recordava das antigas canções em bo. Em novembro de 2009, no mesmo local, também a morte de Boro F. representara o fim da última falante do khora, outra língua andamanesa. Conforme a Profa. Abbi, que mantém o projeto VOGA (Vanishing Voices of the Great Andamaneses), é provável que essas línguas remontem à era pré-neolítica, quando os primeiros homens se instalaram na região, contando elas, portanto, com quase setenta mil anos (cf. Abbi, *Vanishing Voices of the Great Andamaneses*).

LEITURA COMPLEMENTAR

Seis mil línguas: um patrimônio em perigo

Ranka Bjeljac-Babic

Universidade de Poitiers, França

Está condenada a desaparecer em curto prazo a imensa maioria das línguas? Os linguistas estimam que uma língua só pode sobreviver se conta com mais de 100.000 falantes. Pois bem, das cerca de 6.000 línguas que existem hoje em dia no mundo, a metade é falada por menos de 10.000 pessoas e um quarto por menos de 1.000. Apenas umas vinte contam com muitos milhões de falantes.

A morte das línguas não é um fenômeno novo. Desde que se diversificaram, pelo menos 30.000 (alguns falam inclusive de 500.000) nasceram e se extinguíram, em geral sem deixar traços. A essa grande mortalidade corresponde uma duração de vida relativamente breve. Escassos são os idiomas, como o basco, o egípcio, o chinês, o grego, o latim, o persa, o sânscrito, o tamil e alguns outros, que lograram atingir 2.000 anos.

O que é uma novidade, entretanto, é a rapidez com que perecem na atualidade. Voltando no tempo, constatamos que a diminuição da diversidade linguística se acelerou consideravelmente em razão das conquistas coloniais europeias, que eliminaram pelo menos 15% das línguas faladas naquela época. E, se no decorrer dos três últimos séculos, a Europa perdeu umas dez línguas, na Austrália não restam mais que 20 das 250 faladas em fins do século XVIII. No Brasil, 540, ou seja, três quartos das línguas morreram desde que se iniciou a colonização portuguesa em 1530.

O nascimento dos Estados nacionais, cuja unidade territorial estava estreitamente ligada a sua homogeneidade linguística, também foi um fator decisivo de consolidação das línguas adotadas como nacionais e de marginalização das demais. Os governos, em seu marcado empenho por instaurar uma língua oficial na educação, nos meios de comunicação e na administração, procuraram deliberadamente eliminar as línguas minoritárias.

Esse processo de homogeneização linguística foi reforçado com a industrialização e o progresso científico, que impuseram novos modos de comunicação, rápidos, simples e práticos. A diversidade de idiomas foi então considerada como um obstáculo aos intercâmbios e à difusão do saber. O monolinguismo passou a ser um ideal. É assim que, em fins do século XIX, surgiu a ideia de uma língua universal (pensou-se, inclusive, em voltar-se ao latim), o que deu lugar a uma proliferação de línguas artificiais. A primeira delas foi o volapük, sendo o esperanto a que teve êxito mais ressonante e a maior longevidade.

Em tempos mais recentes, a internacionalização dos mercados financeiros, a difusão da informação por meios de comunicação eletrônicos e os demais avatares da mundialização contribuíram para acentuar as ameaças que pesavam sobre as línguas “pequenas”. Uma língua que não está na Internet é uma língua que quase “deixou de existir”. Fica à margem do

“comércio”.

O ritmo de extinção das línguas alcançou, assim, proporções sem precedentes na história: dez a cada ano, em escala mundial. Segundo os prognósticos mais sombrios, de 50% a 90% das línguas faladas hoje em dia morrerão no curso do presente século. Preservá-las é um assunto urgente.

As consequências do desaparecimento das línguas são graves em mais de um sentido. Em primeiro lugar, se nos tornássemos uniformemente monolíngues, é possível que nosso cérebro acabasse afetado ao ponto de perder parte de sua capacidade inata de criação linguística. Na sequência, todas as tentativas de remontar às origens da linguagem humana se tornariam impossíveis e o mistério do “primeiro idioma” jamais seria elucidado. Por último, com a morte de cada língua, um capítulo da história da humanidade se fecha para sempre.

O plurilinguismo é o reflexo mais fiel do multiculturalismo. A eliminação do primeiro acarretará inevitavelmente a perda do segundo. Impor um idioma – seja regional ou internacional – a populações cuja cultura e estilo de vida não se identificam com ele é fazer calar a expressão de seu espírito coletivo. As línguas não só são o meio primordial de comunicação entre os seres humanos, como encarnam também a visão de mundo de seus falantes, sua imaginação, suas formas de transmitir o saber. Apesar de seu parentesco, refletem de maneira diferente a realidade. Se tratarmos de inventariar as diferentes palavras que existem em todos os idiomas para expressar estritamente o mesmo sentido, damos conta de que há no máximo 300, tais como ‘eu’, ‘tu’, ‘nós’, ‘quem’, ‘que’, ‘não’, ‘tudo’, ‘um’, ‘dois’, ‘grande’, ‘comprido’, ‘pequeno’, ‘mulher’, ‘homem’, ‘comer’, ‘ver’, ‘ouvir’, ‘sol’, ‘lua’, ‘estrela’, ‘água’, ‘quente’, ‘frio’, ‘branco’, ‘preto’, ‘noite’, ‘terra’ ...

O perigo que ronda o plurilinguismo é análogo ao que afeta a biodiversidade. Não só porque a grande maioria das línguas são “espécies” em vias de desaparecimento, como também porque entre a diversidade biológica e a diversidade cultural existe um laço intrínseco e causal. Do mesmo modo que as espécies vegetais e animais, as línguas em perigo são endêmicas, ou seja, estão confinadas em regiões exíguas. Mais de 80% dos países onde existe uma “megadiversidade” biológica formam parte dos que albergam o maior número de línguas endêmicas. Essa correlação explica-se pelo fato de que os grupos humanos, ao adaptar-se ao entorno em que evoluem, criam um conhecimento especial de seu ambiente, que reflete em sua língua e, muitas vezes, unicamente nela. Grande parte dos recursos naturais em perigo só são conhecidos atualmente por alguns povos cujas línguas estão em extinção. Ao morrer, estas levam consigo todo o saber tradicional sobre o meio-ambiente.

Em 1992, a Cúpula do Rio criou dispositivos para lutar contra a redução da biodiversidade. Chegou a hora do “Rio das línguas”. A tomada de consciência da necessidade de proteger esse patrimônio surgiu em meados do século XX, quando os direitos linguísticos se integraram à Declaração Universal dos Direitos Humanos. Desde então tiveram início diversos projetos internacionais voltados para salvaguardar o que agora se reconhece como patrimônio da humanidade. Ainda que não logrem pôr término ao processo de extinção das línguas, têm o mérito de atenuá-lo e de promover o plurilinguismo no mundo.

EXERCÍCIO

Nos próximos capítulos vamos estudar as principais famílias linguísticas. A fim de preparar esse novo passo, tome como referência a tabela 1 (as línguas maternas faladas por mais de cinco milhões de pessoas) e verifique quais são as famílias linguísticas representadas por maior número de línguas, completando o quadro abaixo:

Famílias	Número de línguas
Indo-europeia	
Nigero-congolesa	
Austronésia	
Altaica	
Afro-asiática	
Tai-Kadai	
Dravídica	
Austro-Asiática	
Ameríndia	
Sino-tibetana	
Urálica	
Cartveliano	
Cartveliano	
Coreano	
Hmong-Mien	
Japonês	
Nilo-Saariano	

Guia de Leitura Texto 6: As línguas do mundo

- (1) Considerando-se a tabela que cataloga as línguas neste capítulo, explique porque algumas vezes os dados da "estimativa 1" e da "estimativa 2" coincidem.
- (2) Explique qual a razão de falantes do inglês nativos da Austrália serem computados como pertencentes à área linguística europeia no *Tabela 2*.
- (3) Qual a área linguística com menor número de línguas?
- (4) Com relação ao número de falantes, qual a área linguística mais representativa?
- (5) Qual a média de falantes por língua da área europeia? E das áreas africana e asiática?
- (6) Em que situação pode-se considerar que uma determinada língua está condenada ao desaparecimento?
- (7) Qual a principal diferença em realizar um agrupamento de idiomas por área linguística e um agrupamento baseado em um critério de parentesco, isto é, um agrupamento por famílias linguísticas?
- (8) Juntas, duas famílias linguísticas abrangem quase 70% da população mundial. Quais são elas?
- (9) Quais as tendências comuns sinalizadas pelas tabelas 1 e 2?
- (10) Explique o motivo pelo qual o estabelecimento da agricultura foi crucial para a expansão de certas comunidades linguísticas e para o declínio de outras.
- (11) Em que região do mundo surgiu, há sete milhões de anos atrás, a espécie humana – embora não com as características cognitivas que possui atualmente?
- (12) Por que o continente sul-americano, que detém características ambientais similares às da África, não possui grandes mamíferos?
- (13) Quando a agricultura se desenvolveu no Oriente Médio? E nas Américas?
- (14) Por que as comunidades de caçadores-coletores eram constituídas por um número pequeno de indivíduos?
- (15) Atentando-se ao Quadro 1, responda por que passa-se de uma organização igualitária para uma estratificação em classes sociais.
- (16) Por que a escrita é capaz de realizar uma mudança de foco e de cognição nas sociedades em que é empregada?
- (17) Por que, nos primeiros estágios representados no quadro, há apenas uma língua e, no último, várias?
- (18) Explique o motivo pelo qual um viajante que atravessa territórios europeus tem a sensação de viver uma realidade de pluralidade linguística enquanto um viajante que atravessa o território brasileiro tem a sensação de estar em um país monolíngüe.
- (19) Por que locais como a Amazônia, o vale do Mississipi, os Andes e Papua Nova-Guiné, em que houve o surgimento da agricultura, não surgiram civilizações poderosas, como as da China, Europa e Oriente Médio?

- (20) Por que os europeus adquiriram predominância indiscutível nos projetos de expansão marítima e colonização de terras no além-mar enquanto a China, com um potencial militar e uma frota naval superiores, não se lançou em uma aventura similar?
- (21) Levando em consideração a Tabela 5, por que se pode afirmar que nas áreas com menos línguas ameaçadas a grande extinção já ocorreu há tempos?

TEXTO 7

SISTEMAS DE ESCRITA

Aléxia Teles Duchowny

Introdução

A escrita foi inventada há aproximadamente 10.000 anos, o que permitiu que os seres humanos transmitissem e acumulassem com mais eficácia, de geração a geração, descobertas e experiência de seus antecessores. Assim, ao nascermos, podemos ter acesso bastante detalhado ao conhecimento adquirido pelas gerações anteriores. Isso permite progressos impossíveis de existir em sociedades ágrafas, isto é, sem escrita, ou em grupos de outros animais. Os sistemas de escrita não só são sofisticados por si só como derivaram de antecedentes complexos. Provavelmente, todos eles têm como origem alguma escrita pictórica primitiva e todos também são artificiais, ao contrário da linguagem oral, natural aos seres humanos.

A escrita pode ser definida como a representação da linguagem falada por meio de signos gráficos. É a interpretação de uma língua falada, seus elementos, seu ordenamento e as relações desses elementos uns com os outros. Estes sinais materiais visíveis podem variar muito de grupo para grupo humano, mas todos apresentam algumas características gerais: (i) elementos básicos e definidos, os grafemas¹⁹; (ii) normas e convenções arbitrárias compreendidas e compartilhadas por uma comunidade de fala; (iii) meios físicos (escrita no papel, na tela do computador etc) para representar o sistema de escrita, permitindo que seus usuários o interpretem.

Horcades (2004, p. 15) faz a seguinte afirmação:

Letras são como abelhas. Uma abelha sozinha é apenas um inseto irracional. Mas, se observarmos uma colméia com seu funcionamento extremamente complexo, com operárias, soldados, babás, faxineiras, zangões e rainha, veremos que esses insetos primitivos desempenham funções bem determinadas. A abelha não tem inteligência individual, mas a colméia possui inteligência coletiva.

Isso acontece também com as letras. Uma letra sozinha não vale nada. Mas letras juntas formam palavras, e palavras são pensamentos.

¹⁹ Grafema, conforme Ferreira (s/d), é um símbolo gráfico constituído por traços gráficos distintivos que permitem o entendimento visual das palavras na língua escrita. É uma designação mais ampla do que letra, porque também abarca diacríticos, ideogramas e sinais de pontuação.

Os sistemas de escrita são muitos e podem ser classificados de várias maneiras, como você verá em seguida. Para Fischer (2003), essas classificações não são uma tarefa fácil, pois os critérios de cada uma podem variar: tipológico, cronológico, genealógico, geográfico... Além do mais, muitos sistemas de escrita são mistos e há uma gama de empréstimos e inovações que tornam semelhantes sistemas sem relação direta um com o outro. Ferdinand de Saussure, em seu *Curso de linguística geral*, acredita que haveria dois sistemas básicos de escrita, o ideográfico e o chamado fonético. Aqui vai uma proposta de desenvolvimento desses dois tipos básicos.

1 Sistemas ideográficos

Nesses sistemas de escrita, cada grafema ou ideograma²⁰ representa um morfema, isto é, há um único símbolo para uma palavra gramatical completa. A palavra é representada por um sinal único e estranho aos sons de que ela se compõe. Os hieróglifos do Antigo Egito, a escrita chinesa, a escrita dos maias e dos astecas e a cuneiforme são exemplos de sistemas ideográficos, mas apenas o chinês sobreviveu até os dias de hoje. Conheçamos um pouco sobre eles.

1.1 Os hieróglifos egípcios

Os antigos egípcios chamavam a escrita hieroglífica de “fala dos deuses”. As inscrições hieroglíficas egípcias mais antigas datam do século III aEC²¹. Trata-se de um sistema bastante complicado, pois os sinais ora exprimem palavras, ora um som. Como nos alfabetos semíticos - usados em línguas como o hebraico, o aramaico e o árabe -, apenas as consoantes são representadas. Evidentemente, na hora da leitura do texto, as vogais são devidamente pronunciadas. Havia aproximadamente 75 fonogramas biconsonantais e 24 signos monoconsonantais, além dos ideogramas, que são a base do sistema, que representam coisas, ações e abstrações. Por exemplo, a palavra *crocodilo* é a figura de um crocodilo e também representa o som “msh”. Os hieróglifos para gato, *miw*, combinam as figuras de *m*, *i* e *w* com o desenho de um gato.

Os hieróglifos são usualmente gravados em pedra, estando dispostos tanto de cima para baixo como horizontalmente, e tanto da esquerda para a direita como da direita para a esquerda.

²⁰ Ideograma é “um símbolo gráfico ou um desenho que representa um objeto ou uma ideia.” (HOUAISS, 2001, p. 1565).

²¹ Era Comum, equivalente à Era Cristã, mas sem conotação religiosa.

1.2 O chinês

Falada por um quarto da população humana, a língua chinesa tem pelo menos 4.000 anos de existência. Inicialmente, os chineses desenhavam os objetos que queriam representar. Com o passar do tempo, passaram a combinar esses desenhos para representar ideias abstratas. Assim, o símbolo que representa “mulher”, se repetido duas vezes, significa “discussão”; “escutar” é representado por “ouvir” seguido de “porta”; “sincero” ou “verdadeiro” = “homem” + “palavra”.

Atualmente, todo caractere chinês é composto de dois elementos, uma chave ou radical e um elemento fonético. Se por um lado esse novo sistema de escrita, denominado *hsing shen*, aumentou o número de caracteres na escrita chinesa, por outro eliminou muitas ambiguidades. A palavra *k'ó* (rio), por exemplo, é composta por um primeiro elemento que significa *fruta* e que também indica como deve ser a pronúncia final da palavra. Adiciona-se a ele o símbolo determinativo *shui*, que significa *água*. Outros exemplos (<http://www.omniglot.com/writing/chinese.htm>):

kung (trabalho manual) + *hsin* (coração) = *k'ung* (impaciência);

kung (trabalho manual) + *yen* (palavra) = *kung* (luta)

Para não haver confusão entre as dezenas de milhares de caracteres diferentes, o número e a forma de cada traço são feitos com bastante rigor. A escrita tradicional é em linhas verticais, de cima para baixo. Entretanto, atualmente, também encontramos textos na horizontal.

1.3 As escritas americanas pré-colombianas

O sistema de escrita dos maias e dos astecas manifestou-se a partir do século III EC. Até hoje não foi completamente decifrado, em parte porque os missionários espanhóis queimaram a maioria dos manuscritos astecas e praticamente todos os documentos maias. As escritas maias e astecas podem ser chamadas de transicionais por se encontrarem um pouco além da fase ideográfica, apresentando alguma fonetização.

Da língua dos maias, presentes na Guatemala, Yucatán (um dos estados mexicanos) e Baixo México, só foi possível isolar os sinais dos meses e dos dias do calendário. O que se chama de Antigo Império Maia pode ser situado por volta do século I EC. Nessa época, a matemática e a astronomia estavam bastante adiantadas e seu calendário cobria um período de 500 milhões de anos. Sobreviveram apenas três manuscritos, preservados pelo bispo Diego

de Landa, que também destruiu grande número de manuscritos maias. Entretanto, chegaram até o presente grandes colunas monolíticas com caracteres e figuras esculpidas em baixo relevo (estelas), pedras ovais (possíveis altares), esculturas e gravuras de metal, osso e madeira. Segundo fontes espanholas, a escrita maia foi utilizada até fins do século XVII, mas mesmo assim continua por decifrar. Para os maias, a escrita foi inventada pela divindade Itzamná, sendo seu conhecimento limitado aos sacerdotes, seus filhos e alguns senhores (<http://www.famsi.org/mayawriting/index.html>).

Abaixo, alguns elementos do silabário maia:

a	e	i	o	u
ba	be	bi	bo	bu
cha	che	chi	cho	chu
ch'a			ch'o	
ha	he	hi	ho	hu
ka	ke	ki	ko	ku
k'a			k'o	k'u

FIGURA 4 - Alguns caracteres maias

Fonte: <http://www.omniglot.com/writing/mayan.htm>

Nos manuscritos astecas, encontrados no México, há narrativas desenhadas e lendas em caracteres figurados, com frases de sentido vago. A escrita asteca era mais pictural do que a maia e quase todos os seus símbolos eram desenhos de algum tipo. Muitos dos símbolos tinham um valor fonético, sendo utilizados principalmente na escrita de nomes próprios, lugares e divindades. É uma escrita, como a maia, que exige uma descrição oral complementar. Os códices astecas são quase todos relacionados às divindades, aos rituais e à astrologia e até o presente encontram-se apenas parcialmente decifrados. Já foi possível identificar divindades, nomes de pessoas e lugares, e entender certas cerimônias.

1.4 A escrita cuneiforme

O sistema cuneiforme (< latim *cuneus*, “cunha”; *forma*, “forma”) é a escrita mais antiga que se conhece. Apesar de decifrado no século XIX, sua origem continua obscura. Isso porque apresenta muitas semelhanças com outras escritas, como a dos egípcios e dos povos do vale do Indo, situado no subcontinente indiano. Por volta de IV aEC, os sumérios invadem e conquistam o sul da Mesopotâmia, situada no Oriente Médio, entre os rios Eufrates e Tigre. No século XX, mais de mil tabuinhas e fragmentos foram encontrados em Ur, uma cidade da Mesopotâmia, com inscrições em língua sumérica. Inicialmente uma escrita figurativa, torna-se transitória e finalmente ideográfica e fonética.

Com o passar do tempo, os escribas suméricos se deram conta de que os caracteres tornavam-se mais nítidos se impressos na superfície de argila endurecida ao sol do que riscados. Assim, as curvas, círculos e linhas foram substituídos por traços curtos e estreitos. Esses traços eram feitos com um pedaço reto de cana, osso, madeira dura ou metal. A escrita era feita da esquerda para a direita. Para evitar as ambiguidades dos símbolos cuneiformes, os escribas criaram os determinativos, uma classe de sinais colocados antes ou depois das palavras. Não eram pronunciados, mas indicavam a classe geral a que pertencia a palavra.

2 Sistemas fonéticos

Os sistemas fonéticos de escrita visam à reprodução da sucessão de sons de uma palavra. Conforme Martins (2002, p. 40),

Um passo de consequências incalculáveis foi dado quando o homem, na tarefa de fixar e de transmitir o pensamento, percebeu que lhe era possível substituir a imagem visual pela sonora, colocar o som onde até então tinha obstinadamente colocado a figura. Dessa forma, o sinal se libertaria completamente do objeto e a linguagem readquiriria a sua verdadeira natureza, que é oral. Decompondo o som das palavras, o homem percebeu que ele se reduzia a unidades justapostas, mais ou menos independentes uma das outras (enquanto som) e nitidamente diferenciáveis.

Essa revolução dará origem a dois tipos de escrita, a silábica e a alfabética. Vejamos cada uma delas.

2.1 A escrita silábica

Um silabário é um conjunto de símbolos escritos que representam sílabas. Em geral, a um som consonantal segue-se um vocálico. O japonês, o grego micênico, e algumas línguas nativas americanas, como o cherokee, são exemplos de escrita silábica.

Tomemos o japonês moderno como exemplo. Essa língua utiliza três alfabetos, o hiragana, o katakana e o kanji. Assim, por exemplo, em cada um dos alfabetos temos quatro símbolos distintos para *ka, ki, ku, ko; na, ni, nu, ne, no; ha, hi, hu, he, ho...*

2.2 A escrita alfabética

Nessa escrita, cada sinal corresponde a uma letra. Em relação à escrita silábica, a alfabética representa não só maior complexidade ideológica como também maior simplificação técnica.

Não se sabe com certeza que povo ou quais povos teria(m) criado o alfabeto. Em relação ao Ocidente, a partir da Era Comum, a escrita alfabética espalhou-se na Europa graças aos gregos e aos romanos. O desejo de propagação do Cristianismo aos pagãos fez com que os apóstolos constituíssem novos alfabetos tomando por modelo o alfabeto lido por eles mesmos. Assim, os alfabetos gótico e eslavo tiveram como modelo o grego.

A Fenícia é o ponto de partida para as especulações relativas ao surgimento do alfabeto. Mas, de onde os fenícios obtiveram tal invenção? Como e quando apareceu? Supõe-se que o alfabeto fenício tenha chegado à Grécia por volta de 900 aEC. Entretanto, há duas importantes diferenças entre eles: o sistema consonantal e o uso de vogais, ausentes na escrita fenícia. Os mais antigos textos que chegaram até nós, datados do século VIII aEC, já apresentam essas transformações.

A questão da origem dos alfabetos pode ser iniciada pelo alfabeto norte-semítico, considerado como a mais antiga escrita alfabética. Era formada por 22 símbolos consonantais escritos da direita para a esquerda, tendo existido por volta do século II aEC.

O exemplar mais antigo de um abecedário completo foi descoberto na costa da Síria, em Ugarit, em 1929. Trata-se de uma escrita impressa em placas de barro que vai da esquerda para a direita. Ela é a atestação da existência de uma provável escrita alfabética semítica por volta dos séculos XVI ou XV aEC. Os atuais Israel e Síria são, provavelmente, as regiões onde se originou a invenção do alfabeto.

Vamos nos ater, aqui, a quatro alfabetos: o hebraico, o árabe, o grego e o latino.

O alfabeto hebraico originou-se da escrita aramaica. Esta, por sua vez, proveio da fenícia. O que se chama de hebraico quadrado é a escrita empregada durante os primeiros séculos da Era Comum para a cópia de textos sagrados. Como o protótipo fenício, possui 22 consoantes, também usadas como sinais numéricos, e sua direcionalidade é da direita para a esquerda. Para se evitar erros de leituras dos textos sagrados, notam-se com sinais - pontos ou acentos - as vogais, a pronúncia das consoantes e o lugar do acento tônico.

O árabe, como o hebraico, é uma escrita consonântica ainda em uso e escrita da direita para a esquerda. A origem do alfabeto árabe é obscura, apesar de a tradição atribuir sua invenção a um membro da família de Maomé. A primeira inscrição atestada em caracteres árabes é uma inscrição trilingue em grego, siríaco e árabe, datada de 512-513 EC. O alfabeto é

constituído de 28 letras que podem apresentar pontos diacríticos acima ou abaixo delas. O árabe foi adotado por muçulmanos falantes de línguas não semíticas. Assim, por exemplo, há na Espanha literatura em língua espanhola escrita em caracteres árabes, as aljamias. Como o chinês, a escrita árabe é “um elemento de unidade, visto que é o veículo da língua clássica compreendida por todos os letrados, ao passo que a língua falada se fragmentou em múltiplos dialetos”, como esclarece Higounet (2003).

A escrita grega tem importância ímpar para o Ocidente: além de ter sido usada para registrar língua e culturas riquíssimas, foi intermediária entre o alfabeto semítico e o latino e deu origem a numerosas outras escritas. As mais antigas inscrições datam do século VIII aEC. É provável, então, que o empréstimo feito pelos gregos aos fenícios de seu alfabeto e sua adaptação date do século I ou II EC. Essa adaptação foi feita regionalmente, sendo possível encontrar diversos alfabetos locais. Finalmente, no século IV, o alfabeto jônico sobrepujou os demais. Era composto de 24 letras, vogais e consoantes, e ia da esquerda para a direita.

Mas por que o alfabeto grego é o ancestral dos demais alfabetos ocidentais? A inovação em relação ao uso das vogais é a resposta certa. Como em grego são as desinências que indicam função e categoria das palavras, era preciso fixar com precisão a sua posição. Os gregos adaptaram o sistema de notação semítica às particularidades de sua língua: transformaram os sinais representativos das guturais em sinais vocálicos e criaram mais três sinais para as aspiradas.

Os mais antigos testemunhos da escrita latina, que chegaram até nós, datam do fim do século VII ou início do século VI aEC. Não há dúvida de que derivam de um alfabeto grego ocidental, provavelmente advindo da Itália que, por sua vez, teria recebido sua escrita por meio dos etruscos.

No século I aEC, o alfabeto latino se apresenta constituído de 23 letras. Nos séculos II e III, surgem duas novas grafias, a *uncial* e a *nova escrita comum*. Por volta dos séculos VI e VII, o ato de escrever se concentrará nos *scriptoria* eclesiásticos, resultado do desenvolvimento da liturgia e do esfriamento da atividade econômica em geral. No início do século 9, podemos observar o retorno a um tipo comum de escrita, a carolíngia, constituída essencialmente nos *scriptoria* franceses do Reno e de Loire. A reforma educacional feita por Carlos Magno, durante seu reinado (768-814), cria uma demanda de manuscritos que favorecerá o desenvolvimento dessa escrita mais normatizada. Esse imperador restaura antigas escolas e funda novas, na tentativa de reviver o saber clássico, unificando e fortalecendo seu império.

A escrita carolíngia inicia a história da escrita medieval e moderna e dá a Europa ocidental um mesmo tipo de escrita. Apenas no século 13 começa a mudar de forma, sendo substituída pela escrita gótica.

Há outras transformações:

Nos séculos XII e XIII, a arte de escrever, até então essencialmente monástica, se difundiu amplamente no mundo secular e leigo. A renovação intelectual e a criação das universidades provocou uma necessidade sempre crescente de livros. A produção e a venda de manuscritos são a partir de então organizadas nas corporações, com estatutos e privilégios, sob o controle das autoridades acadêmicas. Por outro lado, o desenvolvimento da administração nos grandes Estados ocidentais, o renascimento do direito romano e a extensão do notariado fizeram crescer consideravelmente a necessidade dos escritos. Por volta de 1200 se constituem as grandes chancelarias - e, logo depois, a burocracia - pontifícias, francesa e inglesa; no início do século XIII também o notariado ganhou da Itália todo o sul da França. E o impulso da vida econômica, do grande comércio, dos bancos multiplicaram igualmente as razões de escrever (HIGOUNET, 2003, p. 58).

Nos séculos XIV e XV, as palavras passam a ser escritas sem se levantar a pena do papel, fragmentando a escrita gótica em uma grande variedade de tipos por toda a Europa. No Renascimento, surge uma escrita para livros, a humanística. Será entre os tipos de escritas da Idade Média que os primeiros fundidores de tipos gráficos escolherão seus modelos e deles se originam nossas escritas atuais.

3 A escrita hoje e amanhã

Dois elementos são de grande importância para mudanças ocorridas nos sistemas de escrita em geral: a imprensa e a informática.

A imprensa e sua grafia mecânica permitiram a reprodução de uma enorme gama de grafemas sempre idênticos a si mesmos e que pouco mudaram desde então. Há muita polêmica em relação ao nascimento dessa nova técnica, mas autores como Higounet (2003) acreditam que seu responsável foi João Gutemberg, em Mainz, na Alemanha, por volta de 1450.

Ao mesmo tempo em que possibilita grande flexibilidade, a informática também favorece a padronização dos sistemas de escrita. Ao redigir um texto utilizando o programa *Word* ou similar, o usuário tem uma gama enorme de opções para a apresentação do sistema de escrita em questão, na tela ou no papel. Entretanto, se todos utilizam o mesmo programa, tem-se apenas as opções por ele propostas. Quanto mais amplo for o uso de computadores pelas pessoas, maior a tendência à padronização. Muitos de nós, por exemplo, já sentem algum desconforto ao escrever um texto a mão. Esta sensação de estranhamento em relação à escrita manuscrita é consequência, evidentemente, da falta de prática. Como o uso da informática para expressão da escrita é algo recente, ainda não podemos avaliar com grande clareza as suas consequências para os sistemas de escrita em geral.

Conclusão

Aqui termina nossa viagem pela evolução da escrita. Conforme Ong (1998), e como você deve ter comprovado após a leitura do texto, a escrita amplia praticamente de forma ilimitada a potencialidade da linguagem, além de reestruturar o próprio pensamento de seus usuários. Nossa consciência e até mesmo nosso discurso oral são profundamente afetados pela tecnologia da escrita. Se somos letrados, processamos o mundo a partir de uma capacidade estruturada pela escrita e não apenas de nossa capacidade inata e natural. Não devemos nos esquecer, entretanto, que o fato de a escrita gozar de grande importância para a nossa cultura não deve diminuir o valor da língua oral e a necessidade de estudá-la e respeitá-la em toda a sua variedade.

LEITURA COMPLEMENTAR

Veja os seguintes vídeos em português produzidos pelo Ministério da Educação e Cultura do governo brasileiro, com consultoria científica do linguista L. C. Cagliari:

Construção da escrita - Parte 1

<http://www.youtube.com/watch?v=oXoGEHyGQzY>

Construção da escrita - Parte 2

<http://www.youtube.com/watch?v=BAzeoLfQerM&feature=related>

Guia de Leitura texto 7: Sistemas de Escrita

- (1) Qual a provável origem dos sistemas de escrita?
- (2) Como a escrita pode ser definida?
- (3) Quais as características comuns aos vários sistemas de escrita?
- (4) Quais os sistemas básicos de escrita propostos por Ferdinand Saussure?
- (5) Os hieróglifos egípcios possuíam grafemas para representar vogais? Que classe de fonemas era representada?
- (6) Em que suporte físico os hieróglifos são usualmente grafados?
- (7) Qual o tempo estimado de existência da língua chinesa?
- (8) Quais os dois elementos essenciais de todo caractere chinês?
- (9) Quais as vantagens e desvantagens do *Hsing Shen*?
- (10) Na escrita tradicional chinesa, como as palavras são dispostas na página?
- (11) Por que a escrita dos maias e dos astecas pode ser considerada transicional?
- (12) Segundo fontes espanholas, até quando a escrita maia foi utilizada?
- (13) Qual a principal temática dos códices astecas?
- (14) As inscrições em língua suméria encontradas na cidade de Ur possuíam caráter estritamente figurativo?
- (15) O que eram os *determinativos*, empregados pelos escribas sumérios?
- (16) O que é um silabário?
- (17) De quantos alfabetos dispõe o japonês moderno? Quais são eles?
- (18) Quais povos disseminaram a escrita alfabética no Ocidente?
- (19) Qual o exemplo mais antigo de escrita alfabética?
- (20) De qual escrita originou-se o alfabeto hebraico?
- (21) Quais as semelhanças da escrita hebraica com o protótipo fenício?
- (22) Somente os muçulmanos falantes de línguas semíticas adotaram o árabe?
- (23) Qual a importância da escrita grega para o Ocidente?
- (24) Qual a principal inovação do alfabeto grego em relação às escritas semíticas?
- (25) Qual a principal contribuição das reformas educacionais de Carlos Magno para a escrita latina?
- (26) Quais os dois elementos de grande importância para as mudanças ocorridas nos sistemas de escrita em geral?
- (27) Quais as possíveis consequências, para a escrita, do uso cada vez mais amplo dos computadores?

TEXTO 8

AS LÍNGUAS INDO-EUROPEIAS

Júlio César Vitorino

Observe o quadro abaixo, que mostra o nome dos dez primeiros numerais em algumas línguas românicas:

QUADRO 1

Os dez primeiros numerais em algumas línguas românicas

Português	Espanhol	Francês	Italiano
um	<i>un</i>	<i>un</i>	<i>uno</i>
dois	<i>dos</i>	<i>deux</i>	<i>due</i>
três	<i>tres</i>	<i>trois</i>	<i>tre</i>
quatro	<i>cuatro</i>	<i>quatre</i>	<i>quattro</i>
cinco	<i>cinco</i>	<i>cinq</i>	<i>cinque</i>
seis	<i>seis</i>	<i>six</i>	<i>sei</i>
sete	<i>siete</i>	<i>sept</i>	<i>sette</i>
oito	<i>ocho</i>	<i>huit</i>	<i>otto</i>
nove	<i>nueve</i>	<i>neuf</i>	<i>nove</i>
dez	<i>diez</i>	<i>dix</i>	<i>dieci</i>

Comparando-se esses numerais – e mesmo considerando que a pronúncia dessas palavras não corresponde exatamente à escrita –, é fácil constatar que existe uma relativa correspondência entre o termo de cada língua e o das demais. Isso é decorrência do fato de todas essas línguas terem como ancestral comum o latim, do qual derivaram através de processos contínuos de mudança linguística (no caso específico desses numerais, as mudanças foram principalmente de ordem fonética).

No caso das línguas românicas, temos documentação histórica que nos possibilita confirmar a expansão do latim pelos territórios ocupados hoje pelas línguas românicas e corroborar a tese da existência de relações de parentesco entre elas. Contudo, recuando no tempo, esse mesmo tipo de relação pode ser estabelecido também entre o latim e outras línguas faladas por povos com os quais os romanos não tiveram contato, a ponto de gerar influências de uma língua sobre outra.

O quadro abaixo registra o nome dos numerais, de dois a dez, em algumas línguas, antigas e modernas, da Europa e da Ásia:

QUADRO 2

Numerais de dois a dez em algumas línguas

Latim	Grego	Sânscrito	Galês	Gótico	Tocário A	Antigo eslavo	Lituano
<i>duo</i>	<i>duo</i>	<i>dvá(u)</i>	<i>dau</i>	<i>twai</i>	<i>wu</i>	<i>duva</i>	<i>dù</i>
<i>tres</i>	<i>treis</i>	<i>tráyas</i>	<i>tri</i>	<i>threis</i>	<i>tre</i>	<i>trije</i>	<i>trys</i>
<i>quattuor</i>	<i>téttares</i>	<i>catváras</i>	<i>pedwar</i>	<i>fidwor</i>	<i>stwar</i>	<i>cetyre</i>	<i>keturi</i>
<i>quinque</i>	<i>pénte</i>	<i>páñca</i>	<i>pump</i>	<i>fimf</i>	<i>pāñ</i>	<i>pēti</i>	<i>penkì</i>
<i>sex</i>	<i>héks</i>	<i>sát</i>	<i>chwech</i>	<i>saihs</i>	<i>säk</i>	<i>sesti</i>	<i>sesi</i>
<i>septem</i>	<i>heptá</i>	<i>saptá</i>	<i>saith</i>	<i>sibun</i>	<i>spät</i>	<i>sedmi</i>	<i>septyni</i>
<i>octo</i>	<i>októ</i>	<i>astá(u)</i>	<i>wyth</i>	<i>ahtau</i>	<i>okät</i>	<i>osmi</i>	<i>astuoni</i>
<i>nouem</i>	<i>enné(w)a</i>	<i>náva</i>	<i>naw</i>	<i>niun</i>	<i>ñu</i>	<i>devēti</i>	<i>devyni</i>
<i>decem</i>	<i>déka</i>	<i>dása</i>	<i>deg</i>	<i>taihun</i>	<i>säk</i>	<i>desēti</i>	<i>desimt</i>

Ainda que as semelhanças entre as palavras correspondentes a cada número não sejam tão evidentes quanto as observadas no quadro anterior, em um exame mais atento pode ser constatado que geralmente existem elementos comuns nas diversas séries. Por exemplo, nas palavras para 'dois', a maioria tem um *d-* e um *-u-* (ou *-w-*, ou *-v-*); para 'três', todas apresentam uma sequência de dental seguida de *-r-*. Além disso, observam-se correspondências sistemáticas, ou seja, assim como as formas para 'dois' e 'dez' começam com *d-* em praticamente todas as línguas apresentadas, em gótico encontramos *twai* e *taihun*, iniciadas com *t-* (o caso do tocário sendo diferente por outras razões); do mesmo modo, a *sex* e *septem*, correspondem palavras iniciadas por *s-* na maioria das línguas, com exceção do grego, cujos correspondentes apresentam *h-* inicial (o galês *chwech* tem uma explicação à parte).

Como você já sabe, o exame minucioso de muitíssimas séries como essas, em diversos campos semânticos, unido à comparação de fatos gramaticais e culturais, levou os estudiosos a postularem a existência de uma língua ancestral comum, da qual teriam se desenvolvido as diversas línguas atestadas. À língua comum dá-se o nome de indo-europeu, enquanto as línguas dela descendentes recebem o nome de línguas indo-europeias.

As línguas indo-europeias dividem-se nos seguintes grupos:

1 Línguas célticas

No primeiro milênio a.C., os celtas ocupavam uma área muito vasta, que ia do oriente da Europa ao extremo ocidente. Os celtas britânicos habitavam o sul da Grã Bretanha; os

celtíberos, grande parte da Península Ibérica; os gauleses, uma extensa faixa do Oceano Atlântico ao Mar Negro; enfim os gálatas, o centro da atual Turquia (Anatólia).

Os do primeiro grupo, cujas línguas formam o céltico insular, se espalharam pelas ilhas britânicas e deram origem a todas as línguas célticas ainda em uso, enquanto as línguas dos demais, que formam o céltico continental, desapareceram, deixando poucos vestígios na onomástica, na toponímia e em algumas inscrições não muito consistentes. As línguas continentais, todas extintas, são o gaúlês, o lepôntico (no norte da Itália) e o celtibérico. As inscrições são em alfabeto grego ou nos diversos alfabetos itálicos: as mais antigas, em lepôntico, remontam ao século VI a.C.; o celtibérico usa a escrita dos iberos, um povo não indo-europeu, cujo alfabeto, parcialmente um silabário, parece ser uma mistura dos sistemas grego e fenício, com modificações.

O celta insular divide-se em dois subgrupos: goidélico e britônico. O goidélico, também chamado «céltico Q», inclui o antigo irlandês e seus descendentes: irlandês (gaélico da Irlanda), gaélico da Escócia, e manxês. As inscrições mais antigas do irlandês primitivo, em alfabeto ogâmico, cuja origem é incerta, vão do século IV ao VII, enquanto a mais antiga literatura, em antigo irlandês, começa por volta do início do século VII. A partir da cristianização, no século V, passa-se a usar o alfabeto latino. O irlandês moderno é configurado pelos bardos no século XIII. No século XVI, com a decadência dos bardos, a língua se diversifica em diversos dialetos regionais. Com o domínio do inglês, hoje o irlandês é falado por menos de 70 mil pessoas, muitas das quais como segunda língua, de modo que o seu futuro é incerto.

O gaélico da Escócia provém de uma colônia fundada por imigrantes irlandeses no IV século, sendo falado por menos de 80 mil pessoas, poucas das quais monolíngues. O manxês, da ilha da Manx, tem origens semelhantes: a sua literatura começa no século XVI ou XVII, mas seu uso declina a partir do fim do século XVIII, sendo que o último falante nativo morreu em 1974, seu uso estando hoje limitado ao de segunda língua.

O britônico era a língua falada pelos celtas que habitavam a Grã-Bretanha antes da conquista romana. Hoje, o grupo britônico, ou «céltico P», engloba o galês, o bretão e o córnico. Nos séculos IV e V, as invasões e a expansão dos anglo-saxões levaram os celtas a se concentrarem, por volta do século VII, a oeste, na região do atual País de Gales, e a sul, na região da Cornualha, de onde saíram imigrantes que se transferiram para o noroeste da França, dando origem, em cada uma dessas regiões, ao galês, ao córnico e ao bretão, respectivamente.

As primeiras e breves inscrições conservadas em galês primitivo datam do século VI ao fim do VIII, conservando alguns textos, geralmente em versões modernizadas nos séculos XII e XIII, época considerada do galês médio, com uma rica literatura. O galês moderno tem

início em 1588, com a primeira tradução da Bíblia. O uso da língua declina lentamente, sendo ela hoje falada por meio milhão de pessoas.

Do bretão há poucas atestações até o século XI, a partir de quando se passa a contar com muitas obras literárias, geralmente textos religiosos ou traduções do francês ou do latim. Com a publicação do primeiro dicionário e gramática, em 1659, tem início o bretão moderno. Contudo, não tendo sido nunca a língua de um centro político ou cultural, o bretão jamais passou por algum processo de standardização, sendo composto por um número de dialetos maior que o de qualquer outra língua céltica moderna. O vocabulário é em grande parte tomado do francês e atualmente a língua é falada por cerca de meio milhão de pessoas, a maior parte residente em área rural.

O corno, bastante próximo de bretão, era falado pelos celtas habitantes do sudoeste da Inglaterra. Os primeiros documentos em antigo corno são glosas dos séculos IX e X e um glossário compilado por volta do ano 1100. Do corno médio há cerca de 10 mil linhas de tradução de peças religiosas do inglês, dos séculos XV e XVI. O período do corno tardio vai até 1777, considerado o ano da morte do último falante nativo.

2 Línguas germânicas

Os povos germânicos, na primeira metade do primeiro milênio a.C., viviam no sul da Escandinávia e nas costas dos mares do Norte e Báltico, na região entre as atuais Holanda e Polônia. Contudo, já nos primeiros documentos históricos seu território aparece bem mais estendido para o sul, até as fronteiras do Império Romano, com o qual frequentemente eles entraram em conflito.

O grupo de línguas germânicas é tradicionalmente dividido em três subgrupos:

- (a) o germânico oriental, hoje extinto, que incluía a língua dos godos, dos vândalos, burgúndios e outras tribos;
- (b) o germânico setentrional, que inclui o antigo nórdico e seus descendentes escandinavos modernos;
- (c) o germânico ocidental, com o inglês, o alemão, o holandês etc.

O rúnico, a língua das primeiras inscrições rúnicas, que utilizam o alfabeto chamado *futhork* (a partir do nome das seis primeiras letras: *f, u, th, a, r* e *k*), provavelmente é uma língua setentrional, mas apresenta pontos em comum também com o germânico oriental, levando alguns estudiosos a postularem um grupo nordoriental, que englobaria ambos.

O germânico oriental, representado principalmente pelo gótico, é atestado quase que inteiramente pelo que resta da tradução do Novo Testamento em gótico ocidental, feita pelo bispo Wulfila, a que se acrescentam uns poucos fragmentos de um comentário ao evangelho

de João, algumas inscrições isoladas e palavras preservadas em outros textos. O alfabeto utilizado foi inventado por Wulfila, com base no grego. Em uma língua chamada gótico da Crimeia, provavelmente falada até o século XVIII, temos uma lista de palavras e frases, recolhidas por volta de 1560, que demonstram tratar-se de um dialeto diferente do gótico bíblico. Outras línguas extintas são o vândalo, a língua do povo que deu origem ao nome *Andalusia*, o burgúndio, o gepídico e o rúgio, as quais desapareceram sem deixar vestígios além de alguns topônimos e antropônimos.

O ramo ocidental possui três línguas com atestação suficiente anterior ao ano 1000: o antigo inglês, o antigo saxão e o antigo alto alemão. As duas primeiras são bastante próximas e deviam ser mutuamente inteligíveis, mas diferem significativamente da última.

Os anglos, saxões e jutas habitavam áreas costeiras de onde se encontram a Dinamarca e o norte da Alemanha e, no século V d.C., começaram a migrar para a Grã-Bretanha. As pequenas diferenças entre esses povos, coletivamente chamados anglo-saxões, logo desapareceram, e a língua da nova população é conhecida como antigo inglês ou anglo-saxão. As primeiras atestações são inscrições rúnicas fragmentárias, de pouco valor linguístico, do século IV ou V. Dentre os textos literários, a poesia, na maior parte anônima e de difícil datação, aparece a partir do séc. VIII, enquanto a prosa começa no início do século VII, prosperando muito na segunda metade do século IX, durante o reinado de Alfredo, o Grande, quando a língua, uma variedade de saxão ocidental, é usada como um instrumento literário equiparável ao latim, em uma intensa atividade que se estende até a conquista normanda, em 1066. A partir de então, até cerca de 1500, é o período do inglês médio, cuja literatura supérstite consiste em textos anônimos, religiosos e didáticos em versos. O mais fino poeta inglês antes de Shakespeare, Geoffrey Chaucer, é dessa época. A partir de 1500, o período do inglês moderno é caracterizado, linguisticamente, pela mudança na pronúncia das vogais, ou *Great Vowel Shift*, que transparece pouco na escrita, a qual reflete ainda a pronúncia do inglês médio.

O antigo alto alemão aparece inicialmente em inscrições rúnicas do século VI. Do século VIII provem um grande número de glosas e a primeira literatura, em forma de poemas curtos e textos religiosos. A maior parte do material existente é datada, entretanto, dos séculos IX e X. A língua é preservada em seis dialetos, diferentes na ortografia e em certos aspectos gramaticais. O médio alto alemão tem início por volta de 1100, a partir de quando se desenvolve um alemão *standard*, baseado no bavarês e no alemânico, mas, depois de um certo tempo do florescimento da poesia do médio alto alemão, não havia mais uma língua literária comum e, por isso, a influência mais importante no desenvolvimento do alto alemão moderno foi a tradução da Bíblia feita por Lutero (o Antigo Testamento tendo sido traduzido em 1522 e o Novo Testamento, em 1534), escrita em um estilo fácil, incorporando elementos

do maior número de regiões possível, baseado principalmente no dialeto da Turíngia, região nativa de Lutero. Também é uma língua germânica, com parte de elementos lexicais hebreus, o iídiche, falado por judeus alemães e, posteriormente, devido a migrações, influenciado também pelas línguas eslavas, principalmente o polonês.

O velho saxão era falado até o século XII em uma região entre o Reno e o Elba. Os principais documentos são uma vida de cristo em versos, do início do século IX e uma tradução do Gênesis, de datação mais difícil. A língua é próxima do antigo inglês, do antigo frísio e do antigo baixo francônio. Suas descendentes modernas são as variedades de baixo alemão faladas no norte da Alemanha. O ápice do baixo alemão como língua literária e administrativa foi durante o tempo da Liga Hanseática (séc. XIII-XIV), com centro em Lübeck, cuja variedade de baixo alemão foi, na época, mais importante que o médio alto alemão.

A oeste do antigo saxão era falado o antigo baixo francônio, língua de um grupo de francos ocidentais. É atestado escassamente do século IX ao XII no sudeste da atual Holanda. A variedade falada em Flandres é antecessora do médio holandês, que começa no século XII, e do holandês, do qual deriva, por sua vez, o *afrikaans*, falado na África do Sul. Também provém do dialeto falado em torno de Flandres o flamengo, que é o dialeto holandês falado na Bélgica. Nas costas do norte da Holanda e da Alemanha e nas ilhas da costa do Mar do Norte fala-se o frísio, cujas primeiras atestações seguras são do século XIII.

O germânico setentrional é representado por uma antiga língua, o antigo nórdico, do qual são descendentes as línguas escandinavas modernas. As primeiras inscrições aparecem no século VII. A literatura, que conserva mais que quaisquer outros aspectos da mitologia e do folclore pré-cristãos, é escrita, a partir de meados do século XII, mas que remonta a uma tradição oral muito mais antiga, em um dialeto falado na Islândia, chamado antigo islandês, que muitas vezes é usado em concorrência com o antigo nórdico.

No século IX, exploradores da Noruega se instalam na Islândia e desenvolvem a língua conhecida como islandês, que é muito conservadora, a ponto dos falantes do islandês moderno poderem ler sem muitas dificuldades as sagas em antigo nórdico sem muitas dificuldades, ainda que, no plano fonético, as mudanças terem sido mais extensas. Semelhante ao islandês é o faroês, falado nas ilhas Faroe, cujas inscrições rúnicas conhecidas vão do ano 1000 ao 1500, além de alguns documentos do século XIII. O norueguês, na sua forma antiga, é reconhecido a partir do século XII. O dinamarquês, provavelmente a menos conservadora das línguas escandinavas, é atestado em inscrições a partir do século XI e em manuscritos a partir do século XIII, mais ou menos da mesma época, temos os primeiros documentos em antigo sueco.

3 Línguas itálicas

As línguas itálicas compreendem muitas das línguas indo-europeias da antiga Itália, bem como as descendentes modernas da principal dessas línguas, o latim, que já em época histórica, através do poderio do império romano, fez declinar o uso das outras línguas faladas na península itálica.

O grupo geralmente é dividido em dois sub-grupos: o latino-falisco e o sabélico, também conhecido como osco-umbro. As línguas sabélicas, faladas por povos samnitas, sabinos e outros povos, localizados na parte central da península e, após migrações, também mais ao sul, na Campânia. As principais são o osco, o umbro e o sud-piceno. Em osco restam cerca de 400 inscrições geralmente muito breves, a partir do IV séc. a.C., em umbro, restam principalmente as famosas sete tábulas de Gubbio, do I séc. a.C., além de poucas dezenas de inscrições dos séculos VI e V a.C., as poucas inscrições em sudpiceno vão do VI ao III séc. a.C.

O latino-falisco compreende, como diz o nome, latim e falisco, línguas faladas em meados do primeiro milênio a.C. em uma pequena área no centro oeste da Itália. Do falisco há cerca de trezentas inscrições, do século VII ao II a.C., geralmente muito breves, de modo que sabemos muito pouco sobre essa língua. O latim era uma língua falada originalmente numa pequena área, próxima à foz do Tibre, ao sul da Etrúria (cuja língua não era indo-europeia). As atestações do latim iniciam-se no século VII, sob forma de breves inscrições; os textos literários remanescentes são de alguns séculos posteriores, enquanto a literatura supérstite parte do século III a.C. O período entre as primeiras inscrições é meados do século II a.C. é conhecido como latim arcaico. Segue-se o latim clássico, que vai até o II séc. d.C. e, depois, o latim tardio, até o fim do império romano. Durante esse período, o latim falado nas diversas partes do território romano, ou latim vulgar, começou a se diferenciar nos dialetos que deram origem as línguas românicas.

O alfabeto latino é uma reelaboração do alfabeto etrusco, que não é língua indo-europeia, o qual, por sua vez, é uma reelaboração do alfabeto grego. Há outros alfabetos que eram utilizados na Itália e sua história é semelhante à do alfabeto latino. Nas inscrições mais antigas, antes da dominação romana, aparece também o alfabeto grego.

Das variedades descendentes do latim vulgar, não há textos sobreviventes anteriores ao século IX. O francês é a primeira a ser atestada e até cerca de 1400 é conhecida como antigo francês. Do dialeto falado nas cercanias de Paris, que se tornou dominante a partir dos séculos XII e XIII, desenvolveu-se o francês standard moderno. No extremo norte, o francês normando, que se espalhou pela Inglaterra após a conquista normanda em 1066, dando origem ao chamado anglo-normando, que floresceu até o advento do médio inglês, no séc. XIV. No sul, desenvolveu-se uma variedade chamada *langue d'oc* (em oposição à *langue d'oïl*),

ou ocitano, do qual há ainda alguns dialetos falados no sul da França, principalmente o provençal, atestado desde o século X. A língua românica com o maior número de falantes hoje é o espanhol, atestado na forma de glosas a partir do século X. O espanhol standard se baseia no castelhano, dialeto do reino de Castela, mas há outras variedades, como o catalão, língua oficial do reino de Aragão, no nordeste da península ibérica, e, no sul, um dialeto hoje extinto, era constituído pelo moçarábico, muito influenciado pela língua dos invasores mouros. No noroeste, o galego é, historicamente, uma variedade de português, hoje a segunda língua românica em número de falantes, cujos primeiros textos importantes remontam aos séculos XII e XIII. A Itália, por sua vez, durante o último milênio foi sempre marcada por uma imensa variedade de dialetos, cujos primeiros testemunhos provêm do século X. O dialeto de base do italiano moderno é o florentino, com influências também do dialeto de Roma. Na Sardenha fala-se o sardo, atestado a partir do século XI, mas que conta com escasso material literário; o dialeto da parte central da ilha (logudorês) é bastante conservador, particularmente no que diz respeito a conservação da pronúncia palatal das velares diante de *e* e *i*. Outra língua importante é o romeno, cujos primeiros textos remontam ao século XVI, isolada no extremo oriental, tem forte influência do húngaro e das línguas eslavas. Outras línguas românicas menores e isoladas se encontram no norte da Itália e na Suíça, tais como o rético e o ladino. Enfim, o dálmata, falado antigamente na costa da Croácia, é hoje uma língua extinta.

4 Grego

O fato mais característico do grego é que, em mais de três milênios de história escrita, os seus dialetos não se transformaram em línguas mutuamente incompreensíveis. O dialeto de mais antiga atestação é o micênico, datados entre os séculos XIV e XII a.C., em uma escrita silábica chamada de “linear B”, encontrada na ilha de Creta e em cidades micênicas do continente. Depois dessas inscrições, por um período de mais de 500 anos, encontra-se apenas uma única inscrição, encontrada em Chipre, datada em torno de 1050.

O alfabeto grego em suas versões mais antigas apresenta muitas variações, mas trata-se de uma adaptação do alfabeto fenício que, como os alfabetos de outras línguas semíticas como o árabe e do hebraico moderno, tem letras apenas para consoantes. A grande inovação dos gregos foi usar certas letras para representar os sons vocálicos, utilizando letras fenícias que representavam sons inexistentes em grego, algumas outras, como o *phi*, *khi*, *psi* e *omega* são adições.

Entre o aparecimento das primeiras inscrições alfabéticas, no século VIII, e o início do século V, coloca-se o chamado período arcaico. No início desse período, fixam-se as formas

dos dois poemas homéricos: a *Ilíada* e a *Odisseia*, ápice de uma longa tradição oral. A forma linguística desses poemas é também a dos chamados hinos homéricos e da poesia de Hesíodo e tem como base o dialeto chamado jônico, misturado a um substrado de eólico. Na poesia lírica desse período, destaca-se o uso do dialeto lébio, usado por poetas como Safo, Alceu, nos séculos VII e VI. A partir de 480 a.C., tem início o período clássico, marcado pela hegemonia de Atenas, que tem seu dialeto, o ático, bastante próximo ao jônico, alçado a categoria de língua literária standard. Todos esses dialetos, ou seja, o ático-jônico, o eólico e o micênico, constituem o subgrupo oriental do grego, que inclui também os dialetos não literários arcado-cipriotas e o panfílio. O outro subgrupo, ocidental, é constituído pelos dialetos chamados dóricos, que também têm importantes usos literários.

Após as conquistas de Alexandre Magno, a língua grega se estende por uma área vastíssima, e tem início o período helenístico, linguisticamente caracterizado por uma variedade simplificada do grego falado conhecida como *koiné*, baseado principalmente no ático, com elementos jônicos e de outros dialetos. Com isso, os dialetos gregos pouco a pouco desaparecem na língua falada, restando apenas uma variedade de dórico, o lacônio, que ainda hoje continua na forma do tsacônio.

Com exceção do tsacônico, no leste do Peloponeso, e dos dialetos gregos da Turquia, todas as variedades do grego moderno são descendentes da *koiné*. Após a independência da Turquia, em 1828, foi criada uma nova língua literária standard chamada *katareúsa*, enquanto que uma variedade de grego chamada demótico se tornou a língua falada standard. Em 1976, o demótico substituiu a *katareúsa* também na língua escrita, mas foi de uma convergência de ambos que surgiu o chamado grego standard moderno.

5 Albanês

O albanês é um conjunto unitário dentro do indo-europeu e se divide em dois dialetos principais. O *geg* é falado no norte da Albânia, em regiões da antiga Iugoslávia (Montenegro, Kosovo, Croácia e Macedônia) e em áreas da Turquia. O *tosk*, no sul da Albânia, Grécia e Itália. De todos os grupos, o albanês foi o último a aparecer em fontes escritas, a partir do século XV. A língua, até o século XVIII é conhecida como antigo albanês. A partir do séc. XIX, fala-se em albanês moderno.

6 Línguas balto-eslavas

O grupo balto-eslavo inclui dois sub-grupos: eslavo e báltico.

As línguas eslavas modernas resultam de diferenciações dialetais iniciadas há cerca de 1500 anos. Ainda hoje, muitas delas são mutuamente inteligíveis, principalmente na forma escrita. Os textos literários têm início em 863, com uma tradução da Bíblia, que não sobreviveu. O alfabeto dessa tradução, não é o que conhecemos como alfabeto cirílico, ainda que Cirilo seja o nome monástico de um dos dois irmãos missionários responsáveis pela tradução, mas sim um mais antigo chamado glagolítico, com letras de origem incerta; o que se chama cirílico, hoje usado para escrever o russo, bielorusso, ucraniano, búlgaro, macedônio, sérvio e várias línguas não eslavas da antiga união soviética (inclusive uma língua românica: o moldavo, nome dado ao romeno usado na Moldávia), é uma adaptação das maiúsculas gregas que surgiu cerca de três séculos depois na Bulgária. Os primeiros textos reservados vêm da Bulgária, em uma língua que logo se espalhou como língua litúrgica entre os eslavos, conhecida como antigo eslavo eclesiástico, ou antigo búlgaro, ou antigo macedônio. Os manuscritos mais antigos são do século X. As línguas eslavas modernas podem ser agrupadas nos subgrupos eslavo oriental, ocidental e meridional.

Até a cristianização, no fim do século X, havia apenas uma língua eslava oriental, pouco diferenciada em relação ao restante do eslavo. A língua dos primeiros textos litúrgicos é chamada de antigo russo, mas na verdade é ancestral, além do russo moderno, também do bielorusso e do ucraniano. O russo moderno tem início no século XVIII. Características do bielorusso, muito próximo ao russo, aparecem em antigos textos a partir do século XIII, mesma data em que se encontram as primeiras manifestações do ucraniano.

Entre as línguas ocidentais, o polonês, com textos a partir do século XIV, tem hoje o maior número de falantes. No norte da Polônia, nas proximidades de Gdansk, fala-se kashubiano, às vezes considerado um dialeto polonês, às vezes uma língua à parte. Duas línguas extintas eram faladas na Polônia, o eslovênio, aparentemente um tipo antigo de kashubiano, e o pomerânio, na costa báltica. Uma terceira língua extinta é o polabiano, conhecido por uns poucos textos e que era falado nas margens do Elba, na atual Alemanha. Ainda hoje, fala-se, na Alemanha, em uma região próxima a Dresden, o sorbiano, com um total de 50 000 falantes, poucos monolíngues; dessa língua, o documento mais antigo é uma tradução do Novo Testamento de meados do século XVI. Mais a sul, fala-se o tcheco, atestado a partir do século XIV, e o eslovaco, que é muito semelhante ao tcheco, do qual começou a divergir a partir do século XV.

As línguas eslavas meridionais começaram a se diferenciar das demais a partir do século VI. Após o antigo eslavo eclesiástico, a primeira língua a ser atestada é o esloveno, com documentos dos séculos X e XI e que só reaparecerá em textos a partir do século XV. A sul, fala-se croata, muito próximo ao sérvio e ao bósnio, mutuamente inteligíveis, às vezes considerados uma única língua, chamada servo-croata. No sudoeste da antiga Iugoslávia,

fala-se macedônio, que só se diferenciou do búlgaro em época bastante recente e cuja literatura começa no séc. XIX. Finalmente, o búlgaro, cujo estágio mais antigo é o antigo eslavo eclesiástico, as divergências começam a aparecer nos documentos a partir dos séculos XII e XIII.

O subgrupo báltico continua hoje com duas línguas, o lituano e o letão, que formam o báltico oriental, enquanto o báltico ocidental, continha o atualmente extinto antigo prussiano. O antigo lituano é atestado em traduções de orações católicas a partir do século XVI, a língua literária moderna aparece a partir do fim do século XIX. Também o letão é atestado a partir do século XVI. A única língua báltica ocidental que se conhece, e só um pouco, é o antigo prussiano, a partir de um vocabulário do século XIV, outro do século XVI e três traduções de catecismos luteranos do mesmo século; a língua se extinguiu por volta do século XVIII. Sabe-se o nome de outras línguas bálticas faladas antigamente, mas não há nenhum documento que as ateste.

7 Línguas anatólicas

As línguas da Anatólia, nome antigo da parte asiática da atual Turquia, só foram reconhecidas como indo-europeias no século XX. A mais preservada é o hitita, do qual se conhecem milhares de placas de argila com escrita cuneiforme da segunda metade do segundo milênio a.C. Outras línguas são o palaico e lúvio cuneiforme, também do segundo milênio a.C., escritos com os mesmos cuneiformes do hitita, o lúvio hieroglífico, do segundo e primeiro milênio, escrito com hieróglifos nativos, e, do primeiro milênio, o lício e o lídio, escritos em um alfabeto derivado do grego. Outras línguas como o cário, o pisídio e o sidético, seguramente indo-europeias, são escassamente documentas.

8 Armênio

Só no fim do século XIX é que ficou provado que o armeno constituía um grupo à parte no domínio indo-europeu, uma vez que é muito grande o número de empréstimos de línguas irânicas; na verdade, do vocabulário herdado restavam no armeno pouco mais de 450 palavras na época das primeiras atestações em inscrições e textos a partir do século V d.C., idade de ouro da literatura armena. A língua dessa época, o armeno clássico, permaneceu como standard literário até o século XIX. O alfabeto, baseado no grego, foi criado especificamente para o armeno, contém 36 letras que representam cada som da língua. Segue-se o armeno médio, língua oficial do reino da Cilícia (séc. XI-XIV). As variantes modernas são o armeno ocidental, uma das duas línguas literárias modernas, baseada no dialeto falado

próximo a Istambul, e o armeno oriental, língua oficial da Armênia, baseado no dialeto falado próximo ao monte Ararat e à capital Erevan. As duas línguas literárias são bastante parecidas, mas ao lado delas existe uma enorme variedade de dialetos, muitos deles bastante influenciados pelas línguas dos vizinhos turcos e caucasianos.

9 Línguas indo-iranianas

As línguas indo-iranianas se dividem em dois subgrupos: índico e iraniano; talvez haja um terceiro, o nuristani. Eram e são ainda faladas não apenas na Índia e no Irã, mas também em uma vasta região da Ásia, que vai do Mar Negro à China. Depois das línguas anatólicas, são as de mais antiga atestação, com os textos mais antigos remontando ao séc. XIV a.C. Ao antigo índico, pertencem dois dialetos: o védico e o sânscrito clássico, muito semelhantes entre si.

O védico é a língua literária da tradição védica, distinguindo-se em védico antigo e védico recente. A diferença não é bem cronológica, pois o védico antigo se baseia em um dialeto ocidental, enquanto no védico recente há um número maior de características dos dialetos centrais. O texto mais antigo é o *Rig Veda* (Veda dos cantos), que reúne mais de mil hinos em dez livros chamados mandalas e remonta ao ano 1000 a.C., mas partindo de uma longa tradição oral precedente. O sânscrito é a língua da literatura clássica da Índia e sua base é um dialeto da Índia central; por isso, o sânscrito compartilha muitas características com o védico recente. O chamado sânscrito clássico é a língua codificada pelo célebre gramático Panini (V-IV séc. a.C.), instrumento utilizado na escolarização, na religião e no discurso literário, ainda hoje usada em certa medida. Os prácritos são línguas literárias da tradição médio-indiana (300 a.C.-200 d.C.), não derivados diretamente do sânscrito, mas de uma tradição paralela; o mais importante dos prácritos antigos é o pali, língua do cânone budista. As línguas modernas da Índia são descendentes dos dialetos que constituem as bases dos prácritos antigos.

As línguas indianas modernas, mais de 200, são faladas por cerca de um quinto da população mundial em uma área dialetal contínua sem divisões nítidas. Na região oriental coloca-se o bengali, falado em Bangladesh e no nordeste da Índia, cuja literatura, das mais antigas entre as línguas indo-irânicas modernas, remonta ao século X ou XI, além de assamês, bihari e oriya. A principal língua da região central é o hindi-urdu, ou hindustani, que representam duas línguas literárias, expressas em alfabetos diferentes, mas que constituem basicamente uma única língua falada. Também são línguas centrais o gujarati e o punjabi, além do romani, a língua dos ciganos, falada hoje principalmente na Europa. No nordeste, a principal língua é o nepali, falada no Nepal e no noroeste da Índia. Na região noroeste,

colocam-se as línguas dárdicas, principalmente o kashimiri, no norte da Índia e do Paquistão, além do sindhi, no sudeste do Paquistão, e do singalês, falado no Sri-Lanka, que foi colonizado por emigrantes vindos do norte, no V séc. a.C. Ao sul, não do sub-continente, mas da área de línguas indianas, fala-se o marathi.

Nas mais antigas inscrições da Índia, encontram-se dois tipos de escrita: brahmi e kharosthi. O primeiro teve uso limitado e desaparece após o V séc. d.C. O brahmi, tem uma história de maior destaque. Em suas origens, parece ser derivado de algum alfabeto semítico, pelo menos em parte; por volta do III séc., ele evolui em dois tipos, um ao norte e outro a sul. Da forma do norte surge o devanagari, escrita do sânscrito e do hindi, que, levada pelos missionários budistas, foi usada pelo tocário, kotanês e tibetano. A forma do sul se desenvolve e dá origem a escrita usada pelas línguas dravídicas no sul da Índia, tais como o telegu, o tamil, kannada e malaio, e também a do singalês. Também uma variedade do tipo sul, por volta do século VI, se espalha pelo sudeste da Ásia, dando origem as escritas modernas da Tailândia, Burma, Camboja e Laos.

As línguas irânicas se dividem, cronologicamente, em antigo iraniano, médio iraniano e línguas irânicas modernas. Em relação ao antigo iraniano, avéstico e antigo persiano são as línguas nas quais existem textos supérstites, mas sabe-se da existência de outras línguas, como o medo e o cita.

O avéstico é a língua do Avesta, uma coleção de textos sacros da religião de Zaratustra. No interior desse *corpus*, fala-se em antigo avéstico, datado como do fim do segundo milênio, e novo avéstico, provavelmente do séc. IX ou VIII a.C. A língua pertence a área oriental do iraniano. A tradição desses textos foi oral, até que no séc. IV se inventou uma escrita alfabética com o propósito de recordar a forma de se recitar os textos avésticos. O antigo persa é a língua das inscrições reais da dinastia Aquemênida do antigo Império Persa, datadas entre os séculos VI e IV; a língua pertence à área sudoeste. Os textos do antigo persa são em uma escrita cuneiforme imitada do cuneiforme mesopotâmico, mas os signos são completamente diferentes. A escrita do cuneiforme persa foi a primeira a ser decifrada, abrindo caminho para que se decifrassem os outros cuneiformes mesopotâmicos, pois muitas das inscrições eram bilíngues ou trilíngues.

No médio iraniano, distinguem-se um subgrupo ocidental e outro oriental, ainda que não tenham correspondência geográfica exata. O período reflete a maior área de distribuição das línguas iranianas, do mar Negro à China. Nada dessas línguas era conhecido até o século XX, quando foram descobertos, principalmente no Turquestão chinês, material em parto, sogdiano, bactriano, kotanês e tumshuquês (além do tocário). De todas essas línguas, apenas o médio persa e o sogdiano parecem ter descendentes ainda em uso. Outras línguas irânicas, como o sarmatiano e o alânico, só se conhecem por testemunhos indiretos e não há nenhuma

literatura preservada. A maioria das línguas médio-irânicas são escritas em formas derivadas do alfabeto aramaico que, como todos os alfabetos semíticos não possui signos para vogais breves e cujos sinais para as longas trazem muita ambiguidade.

No irânico ocidental, o médio persa foi a língua oficial da dinastia Sassânida (224-652), mas é conhecida também pela literatura dos séculos IX e X d.C. Foi a língua culta do maniqueísmo na Pérsia. Os dois tipos principais são o pahlavi, a língua standard dos textos do zoroastrismo, e o a língua dos textos maniqueístas, que é mais pura. Nenhum dos dois parece descender diretamente da língua das inscrições aquemênidas.

No subgrupo oriental, coloca-se o bactriano, conhecido por moedas, uma inscrição de 25 linhas e documentos legais, em alfabeto grego; o kotanês e o tunshuquês, situados na rota da seda, cujos documentos são escritos em alfabeto brahmi, o sogdiano, no atual Uzbequistão, preservado em textos variados, que tem um continuador no moderno yaghnobi, e, finalmente o choresmiano, conhecido principalmente a partir de uma tradução interlinear de uma enciclopédia árabe medieval.

As línguas irânicas modernas são faladas em uma área muito vasta, que vai do Cáucaso (ossético), ao norte de Oman (kumzari) e ao Xinjiang (sarikoli). Entre as línguas mais faladas do irânico ocidental, podem-se citar o persa moderno ou farsi, língua nacional do Irã e uma das duas línguas oficiais do Afeganistão, cuja variedade é chamada dari. Um dialeto do persa, o tajiki, é falado no Tajikistão e países vizinhos; o kurdo é falado no Iraque, Irã, Turquia e Rússia; o baloqui, principalmente no Baloquistão, mas também no sudeste do Irã e sudoeste do Paquistão. No irânico oriental, coloca-se o pashto, a outra língua oficial do Afeganistão, cuja tradição literária remonta ao século XVI. Enfim, isolado das demais línguas irânicas, o ossético é falado na Turquia e Geórgia.

10 Tocário

O tocário só foi descoberto no século XX, e em 1907 já foi reconhecido como língua indo-europeia. A maioria dos textos são traduções, muitas bilíngues, de peças budistas, o que facilitou a sua decifração. Os documentos se colocam entre o séc. VI e VIII, e provém de uma região localizada no Turquestão. São duas línguas: uma ocidental, o tocário A, outra, que se encontra na mesma região e também na parte oriental, é conhecida como tocário B. Sobre os falantes dessas línguas, sabe-se muito pouco, e sua identificação é problemática.

Na sua grande maioria, os textos em tocário são escritos em uma versão modificada do alfabeto Brahmi indiano, também utilizado pelo kotanês (língua médio iraniana) e o tibetano. Há também alguns fragmentos do tocário B que usam a escrita maniqueia.

11 Outras línguas

Algumas línguas indo-europeias são atestadas de forma tão fragmentária, normalmente pequenas inscrições, nomes próprios, topônimos e glosas, que a sua colocação em um ou outro sub-grupo não pode ser feita seguramente; são elas: frígio, trácio, macedônio, ilírio, venético, messápico e lusitano. O antigo frígio é conhecido por menos de cem inscrições que vão do século VIII ao V, em um alfabeto derivado de uma forma antiga do alfabeto grego; muito tempo depois, é atestado o novo frígio, nos séculos I e II d.C., documentado por outra centena de inscrições em alfabeto grego, além disso, há algumas glosas de Hesíquio que explicam, em grego, palavras frígias.

O trácio, cujos falantes, na antiguidade, ocupavam a região da atual Bulgária, é atestado em moedas, a partir do VI séc. a.C., e em poucas e breves inscrições, sem interpretação segura, além de cerca de 80 glosas de Hesíquio. Do macedônio, que não deve ser confundido com o moderno macedônio, uma língua eslava, e que era falado no norte da Grécia na antiguidade, temos apenas glosas, uma vez que as inscrições da Macedônia são em grego, adotado como língua oficial no séc. V a.C. Muito pouco conhecido também é o ilírio, da região balcânica, atestado apenas em topônimos, antropônimos e poucas glosas. O venético, no nordeste da Itália, é conhecido através de cerca de 200 inscrições curtas, do VI ao I séc. a.C., as mais antigas no alfabeto etrusco, as mais recentes no latino. Enfim, no extremo oeste do domínio indo-europeu, na península ibérica, foram encontradas três inscrições em uma língua indo-europeia, escritas com alfabeto latino no século I a.C., o lusitano.

Guia de leitura Texto 8: As Línguas Indo-Europeias

- (1) Por que as palavras do quadro 1 são muito semelhantes entre si?
- (2) Como é possível saber que as línguas românicas têm a mesma língua mãe?
- (3) Em relação ao quadro 2, cite as características mais comuns dos termos cognatos.
Exemplo: Palavra "dois": a maioria das línguas apresenta um *d* no início da palavra e um *u*, *w* ou *v* (que são semelhantes) na segunda posição.
- (4) Quais são os dois fatores que permitem que os estudiosos postulem a existência de uma língua ancestral comum chamada de indo-europeu?
- (5) Cite as duas grandes subdivisões do grupo céltico.
- (6) Qual é a língua do celta insular falada até hoje? Onde é falada e quantos são seus falantes?
- (7) Onde é falado o gaélico, atualmente, e quantos são seus falantes?
- (8) Onde é falado o galês, atualmente, e quantos são seus falantes?
- (9) Onde é falado o bretão, atualmente, e quantos são seus falantes?
- (10) Cite os três subgrupos das línguas germânicas e as línguas que os compõem.
- (11) Qual é o local de origem do grupo germânico e a sua provável datação?
- (12) Qual é a principal documentação na língua gótica?
- (13) Indique a periodização do inglês (nome do período, data, atestações).
- (14) Indique a periodização do alemão (nome do período, data, atestações).
- (15) Complete o quadro abaixo:

País	Língua germânica
África do Sul	
Bélgica	
Islândia	
Ilhas Faroé	
Noruega	
Dinamarca	
Suécia	
Holanda	

- (16) Indique a periodização do latim (nome do período, data, atestações).
- (17) O que é o latim vulgar?
- (18) Qual a origem do alfabeto latino?

(19) Complete o quadro abaixo:

Área	Língua românica
França	
Sul da França	
Espanha	
Catalunha, Espanha	
Noroeste da Espanha	
Itália	
Sardenha, Itália	
Romênia	

(20) Qual é a característica mais marcante do grego?

(21) Indique a periodização do grego (nome do período, data, atestações).

(22) Defina *koiné*. Qual a sua relevância?

(23) Quais são os dois principais dialetos do albanês?

(24) Cite as duas línguas eslavas mais faladas atualmente.

(25) Qual é a principal atestação do hitita?

(26) Por que se demorou tanto a perceber que o armênio era uma língua indo-europeia distinta?

(27) Quais são os subgrupos das línguas indo-iranianas e onde são faladas?

(28) Caracterize o bengali.

(29) Caracterize o hindi.

(30) Quais são as línguas irânicas mais faladas, atualmente? Onde?

(31) Caracterize o tocário.

TEXTO 9

AS LÍNGUAS DA ÁFRICA

Tommaso Raso e Jacyntho Lins Brandão

A África é o único continente em que aparecem cinco dos seis grandes grupos étnicos humanos – brancos, negros, coissãs, pigmeus e asiáticos (o único grupo que lá não se encontra sendo o dos aborígenes australianos) – e também, como você já sabe, lá se concentra cerca de um quarto das línguas do mundo. Por que tanta diversidade? Podem ser apontadas duas razões.

A primeira, porque foi lá que o *homo sapiens* se originou há cerca de sete milhões de anos. Portanto, houve mais tempo para a humanidade diferenciar-se, o que propiciou interações entre os povos como em nenhum outro local e, assim, combinações genéticas maiores.

A segunda razão estaria no fato de que a África apresenta grandes diferenças geoclimáticas. Estendendo-se das regiões temperadas do hemisfério norte às temperadas do hemisfério sul, possui algumas das mais altas montanhas tropicais do planeta, uma floresta impenetrável no equador, um deserto praticamente intransitável que separa a parte norte do resto do continente, além de várias outras áreas isoladas. Nesse contexto, desenvolveram-se grupos muito diferentes graças às diversas condições ambientais.

Em tempos históricos, cumpre ressaltar que, por volta de 1400, a situação da África era a seguinte: os brancos povoavam a região norte-saariana; os negros, grande parte da subsaariana; os pigmeus, a floresta pluvial centro-africana e as áreas em volta; os coissãs, a parte sul do continente; os indonésios, Madagascar.

Dentre estes, os pigmeus eram coletores e os coissãs, que compreendem dois subgrupos, os hotentotes (*khoikhoi*) e os bosquímanos (*san*), respectivamente pastores e coletores. Dos *san* há hoje poucos remanescentes no deserto de Calaari, na divisa entre a África do Sul e a Namíbia, ou seja, na área para a qual foram expulsos e na qual conseguiram sobreviver. Mas houve época em que ocupavam grande parte do sul do continente, tendo deixado pinturas rupestres, as mais antigas contando com 27 mil anos. Um pequeno grupo ainda ocupa uma área limitada da Tanzânia, testemunhando assim a antiga extensão da etnia. Os *khoi* contam hoje com menos representantes que os *san*, pois foram massacrados pelas armas e doenças dos europeus, com os quais boa parte dos sobreviventes se miscigenou.

O caso de Madagascar é intrigante. Essa ilha fica a somente quatrocentos quilômetros

da costa africana, estando separada da Ásia pelo Oceano Índico. Nela há duas populações diferentes: os negros e os asiáticos. A língua falada por todos, o malgaxe, é austronésia, parecida com o ma'anyan, falado em Borneu, a 6.500 quilômetros ao leste, nenhum povo semelhante aos indonésios vivendo numa distância menor que esta. Quando, em 1500, os portugueses chegaram a Madagascar, os indonésios já estavam lá.

1 As famílias linguísticas

A classificação das famílias linguísticas africanas foi feita pelo fundador da tipologia linguística, Joseph Greenberg, que iniciou, em 1945, a publicação de uma série de artigos a esse respeito, depois reunidos em *Studies in African Linguistic Classification*, de 1955, livro a que se somou, em 1960, *The Languages of Africa* (a segunda edição revista apareceu em 1966). Ele distingue, na África continental, quatro grandes famílias, algumas com várias subdivisões, cuja distribuição espacial você poderá observar no mapa a seguir (em que os nomes dos grupos se encontram em espanhol). Observe-se que a organização interna de cada família sofreu, nos últimos anos, algumas pequenas reformulações ou correções, sem que se modificassem as linhas gerais propostas por Greenberg (cf. BONVINI, 2008, p. 22-26):

1. Família afro-asiática (em verde no mapa)

Compreendendo 353 línguas, faladas por mais de duzentos milhões de pessoas no norte da África, e dividida em seis grupos:

- 1.1. Semítico, com o árabe, o hebraico, o amárico (língua oficial da Etiópia), o tigrínio (língua oficial da Eritreia), o aramaico etc;
- 1.2. Egípcio-copta, que compreende o egípcio antigo, documentado de 2600 a 700 a.C., bem como suas continuações, o demótico, cujo uso se estende do século VII a.C. ao século V d.C., e o copta, falado do século V ao XVII e hoje ainda em uso como língua litúrgica das igrejas cristãs do Egito e da Núbia;
- 1.3. Líbico-berbere, com cerca de vinte línguas falada na Argélia e no Marrocos, como o tuaregue e o berbere;
- 1.4. Cuxítico, com em torno de trinta línguas faladas na Etiópia, na Somália e no Quênia, dentre as quais o somáli, oromo, beja e afar;
- 1.5. Chádico, que compreende mais cerca de duzentas línguas faladas de Gana até a República Centro-africana, a maior das quais é o hauçá, com vinte e cinco milhões de falantes;
- 1.6. Omótico, com cerca de vinte línguas faladas na Etiópia e no Quênia).

3. Família nigero-congolesa (denominação que, nos autores mais recentes, substitui a de congo-cordofaniana, proposta por Greenberg)

A mais expandida em termos geográficos (em cinza no mapa), englobando 1495 línguas, faladas por mais de 190 milhões de pessoas e distribuídas em nove grupos:

- 3.1. Cordofaniano (cordofanês);
- 3.2. Atlântico ocidental (uolofe, fulani, serere, diola etc);
- 3.3. Mandê (solinquê, suçu, malinquê, bambara etc);
- 3.4. Voltaico ou gur (senufo, moci, grunce, bariba, kassim);
- 3.5. Kwa (baulê, fon, ewe e o subgrupo gbe);
- 3.6. Kru (grebo, betê);
- 3.7. Ijoide (ijó);
- 3.8. Adamaua-ubanguiana (banda, ingbandi, gbaia)
- 3.9. Benue-congolês, o grupo mais importante em termos de sua extensão geográfica e do número de falantes, o qual, por sua vez, se subdivide em onze subgrupos (dos quais dez se situam principalmente na Nigéria):
 - 3.9.1. Defoide (iorubá, igala);
 - 3.9.2. Edoide (edo, urobo);
 - 3.9.3. Nupoide (nupe, ibira, guári);
 - 3.9.4. Idomoide (idoma, igede)
 - 3.9.5. Iboide (ibo);
 - 3.9.6. Cross-river – na confluência dos rios Níger e Benue (efique, ibíbio, ogoni);
 - 3.9.7. Cainji (cambari);
 - 3.9.8. Platoide, ou línguas do platô (berom);
 - 3.9.9. Tarocoide;
 - 3.9.10. Jucunoide;
 - 3.9.11. Bantoide, subdividido, por sua vez, em dois ramos:
 - 3.9.11.1. Bantoide do norte, com línguas faladas na Nigéria e no oeste de Camarões;
 - 3.9.11.2. Bantoide do sul, com diversos subgrupos (ecoide, tivoide, grassfields etc.), o principal dos quais é o banto, com várias línguas, como o umbundo e o quimbundo.

4. Família coissã (em lilás)

Com 22 línguas faladas na África do Sul e na Namíbia, poucas delas contando com mais de mil falantes, como o kwadi e o sandawe, encontra-se ao longo do deserto de Calaari.

Além das línguas propriamente africanas, naquele continente são faladas também outras duas que pertencem a outras famílias: o Afrikaans, na África do Sul, da família indo-europeia (procedente do holandês); e o malgaxe, em Madagascar, da família austronésia. No que diz respeito ao Hazda e outras línguas apontadas, no mapa, como isoladas, há linguistas que defendem pertencerem ao grupo coissã.

Agora observe como a distribuição espacial dos quatro grupos principais nos permite inferir algo da história linguística da África. Dois deles, o nilo-saariano e o coissã não se apresentam em áreas contínuas, o que leva a supor que as línguas que os integram fossem faladas em partes mais extensas do território africano, posteriormente ocupadas, no norte, por línguas afro-asiáticas, e, nas áreas centrais e do sul, pelas congo-cordofanianas.

Na verdade, Greenberg mostrou que o que sabíamos sobre as línguas da África era muito limitado. Uma ideia comum, por exemplo, era que as línguas semíticas (como o hebraico, o aramaico e o árabe) eram originárias do Oriente Médio, tendo-se mostrado, a partir de então, que elas constituem somente um dos grupos da grande família afro-asiática, os demais estando confinados na África. Ainda hoje, doze das dezenove línguas semíticas existentes são faladas apenas naquele continente.

A outra grande surpresa provocada pelas descobertas de Greenberg diz respeito aos pigmeus. Vivendo numa área isolada da floresta equatorial, tiveram eles tempo suficiente para desenvolver uma família linguística própria, hoje completamente extinta. Esse dado, aliado ao fato de que os grupos de pigmeus são atualmente fragmentários, faz concluir que a expansão congo-cordofaniana deve ter eliminado as línguas e quase exterminou também a etnia dos pigmeus. A distribuição das línguas nilo-saarianas mostra algo parecido e, portanto, é provável que esses povos também tenham sido subjugados pelos falantes de línguas afro-asiáticas e congo-cordofanianas.

As línguas coissãs são caracterizadas por serem as únicas no mundo que possuem sons avulsivos, chamados também de cliques, os quais passaram delas para várias línguas bantas. Todas as línguas coissãs sobreviventes estão no sul do continente, com a exceção de duas, isoladas na Tanzânia, a quase dois mil quilômetros da área maior, o que testemunha que a extensão da família devia ser mais ampla, abrangendo áreas depois dominadas pela família congo-cordofaniana.

Como você observou no mapa, esta última ocupa quase toda a África subsaariana, não

se sabendo ao certo o ponto a partir do qual se expandiu. Greenberg notou que todas as línguas da família faladas ao sul do equador pertencem ao grupo banto, com pouco menos da metade das 1.495 línguas nigero-congolesas e mais da metade de seus falantes. Além disso, um traço marcante é que as línguas bantas são muito parecidas entre si, o que sugere que tenham tido pouco tempo para se diferenciarem.

Quase todas as línguas bantas se encontram concentradas na África ocidental e aquelas mais parecidas com as outras línguas da família nigero-congolesa se localizam na divisa entre Camarões e Nigéria. Isso sugere que pode ser nesse local que elas tiveram origem e que foi a partir daí que conquistaram quase todo o restante do continente, avançando em direção ao sudeste. Essa expansão não pode ser recente, porque a língua originária se dividiu em quinhentas línguas dela derivadas, mas também não pode ser muito antiga, porque a diferenciação não é muito grande.

Quando os europeus, no séc XIV, chegaram ao continente africano, encontraram cinco áreas agrícolas: a África norte-saariana, que sempre mantivera contatos com a Europa e que tinha sido inclusive província do Império Romano; o Sahel, logo abaixo do Saara, no lado ocidental; a Etiópia; a África ocidental, na área de origem dos povos bantos; finalmente, a África tropical oriental, onde se encontravam produtos de origem asiática, cuja procedência, não fosse a situação de Madagascar, seria difícil de explicar. Ora, todas as plantas da agricultura africana foram domesticadas ao norte do Equador e o único animal doméstico originário do continente é a galinha d'Angola. Isso poderia explicar por que os bantos, que eram agricultores, dominaram os pigmeus e os coissãs, caçadores-coletores.

A linguística nos diz muito sobre a história da agricultura africana. No sul da Nigéria, onde se falam línguas nigero-congolesas, os nomes das plantas podem ser classificados em três grupos: o primeiro compreende termos presentes em todas as línguas da grande família, nomeando espécies de origem local, como o inhame africano, a palmeira de óleo e a cola; no segundo, há nomes semelhantes somente no interior dos subgrupos, relativos a vegetais de origem asiática, como as bananas; o terceiro abarca nomes que não podem ser relacionados pela origem linguística. Assim, é possível perceber quais são as culturas autóctones e qual a época da introdução das demais na África. Através da glotocronologia, que estuda a mudança linguística para datar a origem das coisas através das mudanças das palavras, inferimos as línguas faladas pelos povos que domesticaram as várias espécies. As famílias que se impuseram na África parecem ter devido esse poder ao fato de que, no local onde se originaram, havia condições para o desenvolvimento da agricultura.

2 As línguas africanas no Brasil

Vamos observar mais de perto o grupo nigero-congolês, ao qual pertence a quase totalidade das línguas faladas pelos negros trazidos para o Brasil (a única de que se tem notícia que não pertence a essa família sendo o hauçá, língua chádica, da família afro-asiática). Ainda que se trate de tema polêmico saber que línguas africanas foram efetivamente faladas aqui e qual sua influência no português brasileiro, seria descabido imaginar que estiveram de todo ausentes e que quase quatro séculos de contato não tenham tido nenhuma repercussão.

Com relação às línguas efetivamente faladas no Brasil-colônia, contamos com dois registros de excepcional importância. O primeiro, o livro de autoria do jesuíta Pedro Dias, que viveu no Rio de Janeiro, publicado em 1697 e intitulado *Arte da lingoa de Angola*, em que se procede a uma descrição do quimbundo de que se tem notícia. Trata-se de uma obra duplamente esclarecedora: “em primeiro lugar, ela testemunha o emprego corrente e habitual, no século XVII, no Brasil, de uma língua africana, o quimbundo, falado por escravos originários de Angola, numa área geográfica extensa, não limitada apenas ao estado da Bahia. (...) Em segundo lugar, esse texto testemunha o olhar lançado por um ‘falante’ português do século XVII, e culto além do mais, sobre uma língua africana” (BONVINI, 2008, p. 37-38).

O segundo registro se encontra na obra de Antonio da Costa Peixoto, em que se recolhe um vocabulário e se reproduzem diálogos na então denominada “lingoa geral de mina” (mina-jeje), falada em Ouro Preto no século XVIII. A primeira versão data de 1731 – tendo o autor, em 1741, elaborado uma segunda versão à qual deu o título de *Obra nova de lingoa geral de mina*. Neste caso, conforme os estudos empreendidos por Castro, trata-se uma língua do grupo kwa, com predominância do fon, ao qual pertence 80% do vocabulário apresentado (cf. CASTRO, 2002).

2.1 As línguas bantas

O grupo banto inclui a maior parte das línguas africanas faladas, em algum momento, no Brasil. Não que aqui tenham aportado apenas povos dessa etnia, mas este foi o contingente maior, desde o século XVI, o que permitiu um contato com o colonizador português mais contínuo e duradouro, resultando em muitos empréstimos e numa muito debatida (mas supõe-se que considerável) influência no português (cf. CASTRO, 2001, p. 34-43; BONVINI, 2008, p. 32-52).

Vamos fazer um experimento. No quadro abaixo você encontrará palavras de origem banta correntes no português, procedentes do quicongo e do quimbundo, já tão integradas ao nosso vocabulário que é provável que você não só não imagine que tenham origem africana,

como também que tenha dificuldades de se lembrar ou mesmo saber as correspondentes de origem portuguesa. Faça o teste:

QUADRO 1

Termos da língua portuguesa de origem banta: Kk = Quicongo; Kb = Quimbundo

Termo português de origem banta	Étimo	Sinônimo português
Bagunça	Kk. <i>bulugusa</i> , Kb. <i>bulungunza</i> , 'bagunça' ²²	<i>Desordem</i>
Banguela	Kb. <i>(ki)bangala</i> , 'fenda (nos dentes)'	
Bunda	Kk./Kb. <i>mbunda</i> , 'bunda', 'ânus'	
Cachaça	Kk. <i>kisasa</i> , 'bebida fermentada, excitante'	
Cachimbo	Kk. <i>(ka)nsingu</i> ; kb. <i>(ka)nzimu</i> , 'pequeno tição fumengante'	
Caçula	Kk. <i>kasuka</i> /Kb. <i>kasule</i> , 'filho mais novo'	
Camundongo	Kk./Kb. <i>kamindongo</i> , 'rato'	
Capenga	Kk. <i>kiapenga</i> ; Kb. <i>kimpenga</i> , 'torto'	
Carimbo	Kk./Kb. <i>kandimbu</i> / <i>kindimbu</i> , 'marca'	
Cochilar	Kk./Kb. <i>kushila</i> , 'cochilar'	
Dengo	Kk./Kb. <i>ndenge</i> , 'criancice', 'cólera pueril'	
Encabular	Kk. <i>(n)kioula</i> ; Kb. <i>kulebula</i> , 'envergonhar'	
Gangorra	Kk. <i>kangala</i> / <i>kangula</i> , 'gangorra'	
Macaco	Kk. <i>makaaku</i> (plural de <i>kaaku</i>), 'espécie de macaco vermelho e cinza, de rabo muito comprido'	
Maconha	Kk./Kb. <i>makonya</i> / <i>makanya</i> , 'variedade de cânhamo'	
Marimbondo	Kb. <i>(ma)di(m)bondo</i> , 'vespa'	
Molambo	Kk./Kb. <i>mulamba</i> , 'pedaço de pano velho'	
Moleque	Kk./Kb. <i>mi-/mu-/na-leeke</i> , 'moleque'	
Zanga	Kk./Kb. <i>nzannga</i> / <i>nzandu</i> , 'zanga'	
Zonzo	Kk./Kb. <i>(ki)nzanzu</i> , 'zonzo'	

Fonte: CASTRO, 2001, p. 135-358.

A integração completa desses termos na nossa língua mostra-se também pelo fato de que são produtivos, ou seja, admitem derivações de acordo com as regras do português: 'bagunçar', 'desbunde', 'capengar', 'dengoso', 'macaquice', 'maconheiro', 'molecada', 'zangar'

²² A estrutura silábica nas línguas bantas prevê a existência apenas de sílabas abertas ou não-travadas, ou seja, terminadas em vogal (do tipo CV – ou CCV, desde que a primeira consoante seja uma nasal): *bulungunza* pronuncia-se *bu-lu-ngu-nza*; *mbunda*, *mbu-nda*; *kansingu*, *ka-nsi-ngu*; *kamindongo*, *ka-mi-ndo-ngo*; *kimpenga*, *ki-mpe-nga*; *kusamba*, *ku-sa-mba*; *ndenge*, *nde-nge*; etc. Outro traço das línguas bantas é o serem tonais, ou seja, não há acento tônico, mas sílabas com diferenças de tons, em alguns casos, dois (alto e baixo, geralmente representados graficamente pelos acentos agudo e grave, respectivamente), em outros, três (alto, médio, baixo).

etc. (cf. Alkmin e Petter, 2008). Isso para não falar de palavras que designam conteúdos transmitidos à cultura luso-brasileira juntamente com suas denominações e para as quais não haveria sinônimos, como ‘macumba’ (< Kk./Kb. *makuba*, *reza*, *invocação*), ‘quiabo’ (< Kk./Kb. *kingombo* > *kingambo* > *kyambo*, ‘quiabo’) e ‘samba’ (Kk./Kb. *kusamba*, ‘rezar’).

O traço mais característico das línguas bantas é a existência de um sistema lexical dividido no que geralmente se denomina “classes” (mas que seria mais adequado entender como uma autêntica divisão em gêneros), o qual varia de língua para língua, comportando, na reconstituição do protobanto, pelo método comparativo, entre dezenove e vinte e quatro categorias. Atenção, não se trata de classes de palavras (substantivo, adjetivo, verbo) e sim de uma organização do vocabulário em gêneros (seres humanos, plantas e objetos, animais etc), a partir de critérios semânticos e morfológicos, com reflexos na sintaxe. Ressalte-se que a distribuição do léxico português entre palavras masculinas e femininas (ou entre masculino, feminino e neutro, como em inglês e alemão) não está em causa, mas uma forma diferente de ver e de organizar o mundo, própria das culturas bantas.

Antes de conhecer os diversos gêneros, vamos entender como o sistema funciona. Cada palavra, no singular e no plural, recebe o prefixo próprio de seu gênero. Assim, por exemplo, em nyankore, o prefixo *mu-* marca o primeiro gênero ou classe, que inclui os seres humanos, o plural sendo feito com o prefixo *ba-*: *muntu* (*mu-ntu*) significa ‘pessoa’ e *bantu* (*ba-ntu*), ‘pessoas, gente, povo’; *ki-* é o prefixo do segundo gênero/classe, relativo, dentre outros, a vegetais, o plural fazendo-se em *bi-*: *kirabyo* (*ki-rabyo*) quer dizer ‘flor’ e *birabyo* (*bi-rabyo*), ‘flores’; o prefixo *m-* marca tanto o singular quanto o plural do terceiro gênero/classe, que inclui os animais: *mbwa* (*m-bwa*) significando tanto ‘cão’, quanto ‘cães’ (KATAMBA, 2006, p. 102).

No quadro abaixo, você encontrará a reconstituição do sistema no protobanto, com os prefixos e o conteúdo semântico que originalmente deveria corresponder a cada classe, e seu resultado no quicongo e no quimbundo, donde são tomados os exemplos:

QUADRO 2
Sistema lexical e de concordância na família banta

Protobanto ²³		Quicongo		Quimbundo		Classe	Exemplos
Sing.	Plural	Sing.	Plural	Sing.	Plural		
1) *mu-	2) *ba-	mu-	ba-	mu-	a-	*seres humanos	<i>mujana, bajana</i> , 'criança(s)'; <i>tata, batata</i> , 'pai(s)' ²⁴
3) *mu-	4) *mi-	mu-	mi-	mu-	mi-	*vegetais e partes do corpo	<i>muti, miti</i> , 'árvore(s)'; <i>mutue, mitue</i> , 'cabeça(s)'
5) *(d/l)i-	6) *ma-	(di)ri-	ma-	ri-	ma-	*líquidos e sólidos	<i>ditadi, matadi</i> : pedra(s) <i>dite, mate</i> : saliva(s) <i>manzo</i> : casario ²⁵
7) *ki-	8) *bi-	ki-	bi(shi)-	ki-	i(shi)-	*depreciativo	<i>kinzo, inzo</i> , 'casebre(s)'; <i>kintu, bintu</i> , 'coisa sem valor'
9) *n-	10) *n-	n-	m(ji)-	-	ji-	*animais e outras coisas	<i>ngombe, jingombe</i> , 'vacas(s)'; (i)nzo, jinzo, 'casa(s)'
11) *du-		lu-	tu-	lu-	malu-	*singulativo	<i>lunkambu</i> , 'um único fio de cabelo'
12) *ka	13) *tu-	ka-	tu-	ka-	tu-	*diminutivo	<i>kanzo, tunzo</i> , 'casinha(s)'; <i>kabumgu</i> , 'vasilhinha'
14) *bu-		bu-	ma-	u-	ma-	*abstratos	<i>bujana</i> , 'infância'
15) *ku-		ku-		ku-	maku-	*infinitivo verbal	<i>kufua, makufua</i> , 'morrer', 'morte(s)'; <i>kuria, makuria</i> , 'comer', 'comida(s)'
16) *pa-		va(ba)-		ga-		*superessivo	<i>panzo</i> , 'sobre a casa'
17) *ku-		ku-		ku-		*alativo	<i>kunzo</i> , 'para casa'

²³ Das várias propostas de reconstituição dos prefixos de classe no protobanto, a primeira das quais se deve a W. H. I. Bleek (1869), adotamos a de Meeussen, *Bantu Grammatical Reconstructions*, p. 97 (*apud* Katamba, 2006, p. 104, onde as demais propostas também são apresentadas).

²⁴ Nomes de parentesco só têm prefixo no plural.

²⁵ O plural desta classe serve para coletivos: *manzo*, 'conjunto de casas, casario', ainda que a palavra 'casa', (i)nzo, pertença à classe 9/10.

18) * <i>mu-</i>	<i>mu-</i>	<i>mu-</i>	*inessivo	<i>munzo</i> , 'em casa'
19) * <i>pi-</i>	<i>fi(pi)-</i>		*diminutivo excessivo	<i>pinzo</i> , 'casa muito pequena'

Fonte: KATAMBA, 2006, p. 104; CASTRO, 2001, p. 32-33.

Esses prefixos, com conteúdos semânticos reconstituídos na forma indicada para o protobanto, nas línguas atuais se apresentam de modo variado, seu número indo de três (no Kako) a vinte e um (no Gana), havendo mesmo o caso de uma língua, o Komo, que eliminou completamente o sistema de classes. Também se constata transferências de conteúdos entre as classes, fazendo com que as distinções se tenham tornado, em parte, semanticamente imotivadas (como, aliás, acontece com o sistema de gênero em português e em outras línguas). As classes mantêm, contudo, sua função gramatical, determinando as regras de concordância nominal e verbal.

Tomemos o exemplo do quimbundo: nesta língua, a relação de genitivo se faz através de um conectivo cuja base é *-a*, ao qual se junta o prefixo próprio de cada classe/gênero do determinado. Assim, se quisermos dizer 'ovos de galinha', será o determinado *maiaki*, 'ovos' (cujo prefixo é *ma-*, uma vez que se trata de palavra da classe IV) que regerá o conectivo *ma* (< *ma+-a*), o determinante permanecendo inalterado (neste caso, *sanji*, 'galinha', palavra da classe IX, que não apresenta prefixo), constituindo-se então o sintagma nominal *maiaki ma sanji*, 'ovos de galinha'. Caso quiséssemos falar 'pescoço de galinha', então seria *xingu*, 'pescoço', palavra da classe IX, que regeria o conectivo *ia*: *xingu ia sanji*, 'pescoço de galinha'. Se fosse o caso de falar da 'morte da galinha', então *kufua*, 'morte' (classe VIII) exigiria o conectivo *kua*: *kufua kua sanji*. E assim por diante.

No quadro abaixo, você encontrará os prefixos de singular e plural do quimbundo, bem como os conectivos de genitivo próprios de cada caso. Você mesmo deverá completar a lista de exemplos a partir das indicações dadas e do modelo apresentado para a classe I (a numeração das classes/gêneros é a fornecida por Chatelain).²⁶ Vamos sempre construir sintagmas com o significado de 'tal coisa do pai', ou seja, o determinante é 'pai', *tata*, o qual permanecerá sem modificações; os determinados (registrados no espaço da extrema-direita do quadro) são os que regem os conectivos correspondentes ao singular e plural de sua classe.

²⁶ O livro de Chatelain foi publicado em 1888/1889, ou seja, ele descreve o quimbundo como falado em fins do século XIX (o que constitui uma vantagem se o interesse é estudar sua influência no português do Brasil). Naturalmente, há algumas diferenças com relação à forma atual (lembre-se que as línguas estão constantemente em mudança). Isso, contudo, não prejudica o nosso propósito de exemplificar como funciona o sistema de gêneros ou classes de concordância nas línguas bantas. Atente-se também para o fato de que, na tabela, adotamos a numeração das classes (em algarismos romanos) tal qual apresentada por Chatelain. Para o conteúdo de cada classe, utilizamos também o artigo de Bonvini.

Para que nosso sintagma fique mais corrente, vamos usar também o artigo definido, que é invariável em termos de gênero e número: *o*. Mãos à obra:

QUADRO 3
Exemplo das regras de concordância no quimbundo

Substantivos			Pre- fixos	Conec- tivo: pref.+ <i>-a</i>	Exemplos
Classe (Gênero)	Núme- ro				
I	Humanos	Sing.	<i>mu-</i>	<i>ua</i>	<i>O mubika ua tata</i> , 'o escravo do pai'
		Plural	<i>a-</i>	<i>a</i>	<i>O abika a tata</i> , 'os escravos do pai'
II	Plantas e objetos	Sing.	<i>mu-</i>	<i>ua</i>	<i>mulembu</i> , 'dedo'
		Plural	<i>mi-</i>	<i>ia</i>	
III	Objetos fabricados, inanimados	Sing.	<i>ki-</i>	<i>kia</i>	<i>kiala</i> , 'unha'
		Plural	<i>ia-</i>	<i>ia</i>	
IV	Corpos, coletivos, líquidos, plantas	Sing.	<i>ri-</i>	<i>ria</i>	<i>rilonga</i> , 'prato'
		Plural	<i>ma-</i>	<i>ma</i>	
V	Abstratos	Sing.	<i>u-</i>	<i>ua</i>	<i>uhaxi</i> , 'doença'
		Plural	<i>mau-</i>	<i>ma</i>	
VI	Objetos	Sing.	<i>lu-</i>	<i>lua</i>	<i>lumuenu</i> , 'espelho'
		Plural	<i>malu-</i>	<i>ma</i>	
VII	Inespecífico	Sing.	<i>tu-</i>	<i>tua</i>	<i>tujola</i> , 'tesoura'
		Plural	<i>matu-</i>	<i>ma</i>	
VIII	Infinitivos e deverbais	Sing.	<i>ku-</i>	<i>kua</i>	<i>kunua</i> , 'bebida'
		Plural	<i>maku-</i>	<i>ma</i>	
IX	Animais, fenômenos naturais	Sing.	∅	<i>ia</i>	<i>kabalu</i> , 'cavalo' ²⁷
		Plural	<i>ji-</i>	<i>ja</i>	
X	Diminutivos	Sing.	<i>ka-</i>	<i>ka</i>	<i>kanzo</i> , 'casinha'
		Plural	<i>tu-</i>	<i>tua</i>	

Fonte: CHATELAIN, 1888-89, p. 1-14; para o sentido de cada classe, BONVINI, 1996, p. 80.

²⁷ É nesta classe que, em geral, se incluem os empréstimos tomados de línguas estrangeiras: *kabalu*, por exemplo, procede do português 'cavalo'. O fato de que, no singular, o que caracteriza a classe IX é a ausência de prefixo facilita a assimilação. No plural, as palavras tomadas de empréstimo se flexionam normalmente: *jikavalu*, 'cavalos'.

Estes prefixos se usam, além de nos sintagmas de genitivo, também nos adjetivos, pronomes possessivos e demonstrativos, bem como nos verbos em terceira pessoa e até nos numerais, constituindo um sistema de concordância que permite o estabelecimento de relações sintáticas bastante coesas. Apenas para ilustrar com mais um exemplo: ‘aquele homem ama seus cães’ diz-se assim (na legenda se indicam a classe dos prefixos e seu número):

<i>riala</i>	<i>riná</i>	<i>rizola</i>	<i>jimbua</i>	<i>jê.</i>
ri-ala	ri-na	ri-zola	ji-mbua	jê (< ji-+é)
CL4 SING-homem	CL4 SING-aquele	CL4 SING-ama	CL9 PL-cão	CL9 PL - dele

Observe como tanto o demonstrativo *riná*, ‘aquele’, quanto o verbo *rizola*, ‘ama’, concordam com *riala*, ‘homem’; por outro lado, repare também que o possessivo *jê*, concorda com o nome que determina, ou seja, *jimbua*, ‘cães’. Caso se queira dizer que ‘aqueles homens amam seus cães’, então o demonstrativo e o verbo passarão a concordar com o substantivo ‘homens’ no plural, *mala* (< *ma-ala*), recebendo o mesmo prefixo de classe, ou seja, *ma-*:

<i>mala</i>	<i>maná</i>	<i>mazola</i>	<i>jimbua</i>	jê
CL4 PL-homem	CL4 PL-aquele	CL4 PL-ama	CL9 PL-cão	CL9 PL - dele

2.2 As línguas do grupo kwa e o iorubá

Este grupo da família nigero-congolesa é constituído por um grande número de línguas tipologicamente muito diferenciadas, faladas no Senegal, Gâmbia, Guiné-Bissau, Guiné Conacri, Serra Leoa, Libéria, Burquina-Fasso, Costa do Marfim, Gana, Togo, Benin e Nigéria, todos países localizados no extremo ocidental da África. Ao Brasil aportaram principalmente falantes de línguas do grupo ewe-fon e de iorubá, assim distribuídos: os jeje-minas, a partir do século XVII, com presença em Minas Gerais, Bahia, Pernambuco, Maranhão e Rio de Janeiro; os nagô-iorubás, a partir do século XVIII, com presença registrada no Rio de Janeiro e nas cidades de Salvador, Recife e São Luís (CASTRO, 2001, p. 39).

O contato dessas línguas com o português no Brasil foi menos prolongado e regionalmente mais localizado que o das línguas bantas, mas nem por isso menos marcante. Os próprios itens lexicais delas recebidos pelo português mostram, em geral, esse caráter regional, cultural e socialmente mais definido, podendo-se citar como exemplos: ‘acarajé’ (Yor. *àkàrà jε*), ‘axé’ (Fon (*na*) *atšè*/Yor. *àšè*), ‘babalorixá’ (Yor. *babalórišà*), ‘gogó’ (Fon *kògò*/Yor. *gògò ngò*), ‘odara’ (Yor. *òdàrá*), ‘orixá’ (Yor. *òrìšà*), ‘Oxalá’ (Yor. *Òšálá*) etc. (ver CASTRO, 2001, p. 135-358).

Sem dúvida a língua que mais se conservou no Brasil foi o iorubá, a qual, na Bahia, “não é mais falada nas ruas, entre as pessoas (...), mas é ainda muito falada nos terreiros dos candomblés (...), nos cânticos para os orixás e durante as festas populares, como o carnaval, a festa de Iemanjá, a festa de São Cosme e de São Damião etc.”, recebendo, em geral, a denominação de “nagô” – podendo-se afirmar que “o nagô é o iorubá brasileiro”, ou “um dialeto do iorubá que se distanciou do iorubá padrão por causa da influência da língua portuguesa” (AJAYI, 2002, p. 304-305).

Das características principais do iorubá, que permitem um contraste com as categorias das línguas bantas (em especial do quimbundo, que nos serviu de exemplo), podem ser arroladas:

- (a) Não há diferenças morfológicas de gênero, as palavras que designam seres humanos e animais opondo-se enquanto itens lexicais para nomear os machos e as fêmeas:

<i>okùnrin</i> , ‘homem’	<i>obinrin</i> , ‘mulher’
<i>bàbá</i> , ‘pai’	<i>iyá</i> , ‘mãe’
<i>oba</i> , ‘rei’	<i>ayaba</i> , ‘rainha’
<i>àkúko</i> , ‘galo’	<i>adiè</i> , ‘galinha’

Os termos que designam ‘homem’ e ‘mulher’ podem aglutinar-se a outras palavras para opor masculinos e femininos, quando se trata de seres humanos: *omokùnrin*, ‘filho’ x *omobirin*, ‘filha’; *omodékòrin*, ‘menino’ x *omodébinrin*, ‘menina’; *erúkunrin*, ‘escravo’ x *erúbirin*, ‘escrava’. Quando se trata de animais, a distinção de sexo se faz com a anteposição dos termos *ako*, ‘macho’, e *abo*, ‘fêmea’: *ako esin*, ‘cavalo’ x *abo esin*, ‘égua’; *ako malu*, ‘touro’ x *abo malu*, ‘vaca’; *ako ajá*, ‘cão’ x *abo ajá*, ‘cadela’.

- (b) Não existem também diferenças morfológicas de número, o plural sendo indicado pela anteposição aos substantivos da palavra *awon* (*òbe*, ‘faca’ – *awon òbe*, ‘facas’) ou pela posposição aos mesmos de um numeral (*eiyele méta*, ‘pombos três’ – ‘três pombos’). Quando se trata de uma enumeração de plurais, *awon* se antepõe apenas ao primeiro termo:

<i>Mo</i>	<i>fe</i>	<i>rà</i>	<i>awon</i>	<i>adiè</i>	<i>pépeíye</i>	<i>ati</i>	<i>àkúko</i> .
Eu	quero	comprar	PL	galinha,	pato	e	galo.

Eu quero comprar galinhas, patos e galos.

- (c) O verbo não se flexiona, as pessoas sendo marcadas pelos pronomes:

emi nri, ‘eu vejo’
ìwo nri, ‘tu vês’

òun nri, 'ele vê'
àwa nri, 'nós vemos'
eyin nri, 'vós vedes'
àwon nri, 'eles veem'.

Também os vários tempos e modos são marcados por elementos antepostos ao verbo: *emi yio rí*, 'eu verei'; *emi iba rí*, eu veria; *emi yio ti rí*, 'eu já vi'; *emi iba ti rí*, 'eu teria visto' etc.

- (d) A construção com genitivo se faz com a anteposição ao determinante da preposição *ti*: *ilé ti bàbá*, 'casa do pai'; *imú ti iyá*, 'nariz da mãe'. (cf. AJAYI, 2002, p. 109-143; CROWTHER, 1852, p. 9-30)

2.3 As línguas nigero-congolesas e o português do Brasil

A influência das línguas africanas no português do Brasil é tema de intenso debate na atualidade, em vista de cinco questões básicas: (a) a escassa documentação sobre a procedência dos escravos que aportaram no país durante quase quatro séculos; (b) os poucos registros sobre as línguas por eles faladas no Brasil, com a exceção dos trabalhos de Pedro Dias e Antônio da Costa Peixoto; (c) o fato de que os traficantes e senhores tinham por hábito mesclar num mesmo espaço escravos de diferentes procedências e, portanto, falantes de diversas línguas, a fim de evitar rebeliões, o que pode ter dado origem a falares crioulos de base portuguesa, todavia não documentados; (d) a dificuldade em estabelecer se os empréstimos e outros influxos se deram no Brasil ou já nos próprios países africanos também colonizados por Portugal; (e) finalmente, a dificuldade em se estabelecer se os rumos do português brasileiro se deve ao contato com as línguas africanas (e ameríndias) ou à própria deriva do português, que em muitos aspectos não difere da que se observa em outras línguas românicas (ver FIORIN; PETTER, 2008).

Em que pesem essas dificuldades, vale repetir, não parece razoável supor que séculos de contato linguístico não tivessem consequências. A existência, documentada no século XX, de línguas cultuais e secretas aponta para a possibilidade de um fenômeno que pode ter tido abrangência mais geral. As línguas cultuais foram empregadas, desde o século XIX, nos cultos afro-brasileiros, sendo transmitidas aos iniciados e dividindo-se em dois tipos principais: (a) as relacionadas com o candomblé e suas divisões internas (o iorubá, na maior parte dos casos; o ewe-fon, na tradição jeje; o quimbundo-quicongo, na angola; etc), constituindo mais formas pidginizadas que propriamente línguas, ou seja, embora se constate a existência de um fundo lexical procedente da língua de referência, não se observa o funcionamento gramatical a ela correspondente; (b) as utilizadas na umbanda, muito próximas do português dito popular, com um vocabulário, um semantismo e marcas morfossintáticas próprias. Já as línguas

secretas são utilizadas por populações negras isoladas, tendo sido documentadas nos dois casos já referidos: na Tabatinga, bairro de Bom Despacho (MG), por Queiroz (*Pé preto no barro branco*); e no Cafundó, em Salto de Pirapora (SP), por Vogt e Fry (*A África no Brasil*). Em ambos os exemplos, constituem “uma espécie de código secreto que serve, de preferência, como meio de ocultar conversas, principalmente na presença de estranhos”, sendo “provável que essas línguas tenham exercido a mesma função no passado, a fim de esconder dos senhores as palavras trocadas entre os escravos, sobretudo quando eles planejavam fugas”, como, aliás, afirma Peixoto em seu trabalho. Em termos linguísticos, “essas línguas partilham um fundo lexical do tipo *banto*, mas enquanto a do Cafundó se aproxima de uma forma pidginizada, a de Tabatinga avizinha-se mais daquilo que se designa como ‘português popular brasileiro’” (BONVINI, 2008, p. 51-52).

Já vimos como há muitíssimos itens lexicais tomados de empréstimo pelo português às línguas africanas, fato sobre o qual não pairam dúvidas: no levantamento de Castro, eles somam 3517 itens, dos quais 1322 são de origem banta (quicongo, quimbundo e umbundo), 1299 de origem oeste-africana (iorubá, fon), 3 de origem imprecisa (banta ou oeste-africana), outros 853 sendo classificados na categoria de decalques (como, por exemplo, ‘o-de-comer’, cf. Castro, 2001, p. 135-358).

Todavia, quando se ultrapassa o nível lexical, torna-se muito difícil determinar o quanto outras tendências do português brasileiro, nos planos fonético, morfológico e sintático, podem dever-se ao influxo dessas línguas. De qualquer forma, temos de considerar a hipótese de que o contato com as línguas africanas pode ter acelerado tendências latentes no português, ou seja, não se trata de o português do Brasil ter adquirido traços alienígenas, mas de as línguas africanas (e também indígenas) terem exercido um papel nos rumos que ele tomou, processo que Castro chama, no campo específico da fonética, de “imantação”.

Abaixo você encontrará um breve resumo de algumas tendências do português brasileiro que se acredita podem ser devidas ao contato com as línguas africanas, em especial das bantas:

1. Como já observamos, a sílaba, nas línguas bantas é de tipo aberto, ou seja, sempre terminam em vogal (CV ou CCV). Desse modo, a divisão silábica das palavras abaixo (as quais você já conhece) se faz como indicado, o que, aliás, torna mais evidente o prefixo de classe:

<i>bantu</i>	–	<i>ba-ntu</i> ;
<i>kanzo</i>	–	<i>ka-nzo</i> ;
<i>rilonga</i>	–	<i>ri-lo-nga</i> ;
<i>mulembu</i>	–	<i>mu-le-mbu</i> .

Isso dá um relevo especial à vogal, o que Castro acredita ter relação com a diferença principal entre o português do Brasil e o de Portugal: enquanto nesta última variedade as vogais átonas se abreviaram ou emudeceram, na nossa elas se mantiveram como o elemento principal da sílaba.

2. Você já sabe que no sistema de classes o número (singular/plural) é marcado por prefixos (e não por sufixos, como em português). Isso pode estar relacionado com a tendência do português brasileiro de deslocar as marcas de plural para a esquerda, em construções como 'os menino', 'as casa'. Não se pode afirmar que o movimento no sentido de simplificação morfológica, conjugado com essa tendência para a esquerda, não seja próprio das línguas românicas (o francês, por exemplo, foi a que tirou mais consequências disso, transferindo a marcação de singular/plural para a esquerda, ainda que a grafia continue a manter a oposição -Ø/-s à direita: singular *l'ami* [la'mi], 'o amigo'; plural *les amis* [leza'mi]). Todavia, poderíamos estar diante de uma conjunção de fatores, em que o contato com as línguas bantas pode ter ativado uma tendência latente no português.

3. Também o verbo nas línguas bantas recebe a marcação de pessoas apenas à esquerda: em quimbundo, *kubanga*, 'fazer', conjuga-se assim no chamado presente futuro (correspondente a nosso presente simples):

eme ngibanga, 'eu faço'

eie ubanga, 'tu fazes'

muene ubanga, 'ele faz'

eta tubanga, 'nós fazemos'

enu nubanga, 'vós fazeis'

ene abanga, 'eles fazem'.

Além dos pronomes pessoais (*eme*, *eie* etc), preste atenção para o fato de que o verbo se flexiona recebendo prefixos de número e pessoa (*ngi-*, *u-* etc) – e não sufixos, como em português –, o mesmo valendo para as desinências de tempo etc. Do mesmo modo que com relação ao item anterior, não se pode dizer a transferência das marcas de flexão verbal para a esquerda seja alienígena às línguas românicas (de novo o francês, dentre outras, serviria de exemplo, cf. Quint, A realização do sujeito em português do Brasil); todavia, usos como 'tu vai' [tu'vai], 'nós vai' [nos'vai], 'eles vai' [es'vai], podem ter sido potencializados pelo contato com as línguas africanas.

4. Você viu que a classe 15 do protobanto é a que comporta os infinitivos verbais, como nos exemplos do quimbundo: *kufua*, ‘morrer’; *kuzola*, ‘amar’; *kuria*, ‘comer’; etc. Ora, os infinitivos podem ser usados também como nomes, a saber: *kufua*, ‘o morrer’ ou ‘morte’; *kuzola*, ‘o amar’ ou ‘amor’; *kuria*, ‘o comer’ ou ‘comida’. Parece que alguns usos do português brasileiro foram decalcados dessa utilização nominal de infinitivos, como em ‘o comer está pronto’. Ressalte-se que no quimbundo a nominalização chegou ao ponto de terem sido criados plurais para os substantivos da classe VIII (com o acréscimo ao prefixo *ku-*, próprio da classe, do prefixo *ma-*, que, da classe 6 do protobanto foi estendido às que, em princípio, careciam de marca de plural, ou seja, as classes 11, 14 e 15).
5. Já chamamos a atenção para o fato de que, no quimbundo, os gêneros expressos pelas classes não incluem a ideia de masculino ou feminino: *mubika*, por exemplo, pode significar tanto ‘escravo’, quanto ‘escrava’. Com poucas exceções (como *riala*, ‘homem’; *muhatu*, ‘mulher’; *tata*, pai; *mama*, ‘mãe’), todos os substantivos são epicenos. Quando se deseja estabelecer a diferença entre macho e fêmea, acrescenta-se à palavra uma determinação (utilizando-se o conectivo próprio da classe) com as palavras *riala* (‘homem’) ou *muhatu* (‘mulher’):

<i>-a riala</i>	<i>-a muhatu</i>
<i>mubika ua riala</i> , ‘escravo’	<i>mubika ua muhatu</i> , ‘escrava’
<i>ribengu ria riala</i> , ‘rato’	<i>ribengu ria muhatu</i> , ‘ratazana’
<i>mulambi ua riala</i> , ‘cozinheiro’	<i>mulambi ua muhatu</i> , ‘cozinheira’
<i>hoji ia riala</i> , ‘leão’	<i>hoji ia muhatu</i> , ‘leoa’
<i>mona ua riala</i> , ‘filho’	<i>mona ua muhatu</i> , filha

Como em outros casos, esse tipo de estrutura parece ter sido decalcado por construções do português brasileiro do tipo ‘filho homem’ x ‘filha mulher’ (‘filho homem dá menos trabalho que filha mulher’); ‘menino homem’ x ‘menina mulher’ (‘menina mulher é mais quieta que menino homem’); etc.

6. É significativo que para Pedro Dias, que escreve no século XVII, a dupla negação seja sublinhada como algo típico da “língua de Angola”: conforme suas próprias palavras, desde que ao verbo se acrescenta a “palavra *cana*, antes ou depois do verbo, fica negativo, v.g. *canangazóla*, não amo, *canángagiba*, não matei. Porém, posta antes e depois do verbo, nega com eficácia, v.g. *canángagiba cana*, não matei não” (*Arte da lingoa de Angola*, p. 21). Poderia a tendência do português brasileiro para a

dupla negação (“não chore ainda não...”) ser devida ao contato com as línguas da África, ainda que o mesmo fenômeno se observe também em línguas românicas como o francês (*je ne sais pas*, ‘eu não sei’)?

7. Uma segunda observação de Pedro Dias aponta também como típica da “língua de Angola” uma tendência que se firmou bastante no português do Brasil, levando a um uso bastante restrito do advérbio de afirmação ‘sim’: “não têm os ambudos casos, e por isso respondem pela mesma pessoa e proposições pelas quais se faz a pergunta, v.g. *Nzambi üazola atu osso? Deus ama a todos? üazola: ama*” (*Arte da língua de Angola*, p. 41). Note-se que esse uso é o mais comum no português falado do Brasil: ‘– Você vai na minha casa? – Vou.’; ‘– Posso pegar esse livro emprestado? – Pode.’

LEITURA COMPLEMENTAR

As línguas africanas no Brasil

Yeda Pessoa de Castro

(In: CASTRO, 2001, p. 62-77)

A questão fundamental (...) é como precisar, no emaranhado de línguas existentes na África, quais dentre elas foram faladas por cinco a oito milhões de indivíduos trazidos para o Brasil por mais de três séculos consecutivos, se a documentação histórica referente ao tráfico e os raros testemunhos que ficaram da época quanto ao modo como vivia a massa escrava não levam em consideração a variedade étnica do negro?

Restam, como ponto de partida, as evidências linguísticas para complementar a lacuna da informação histórica subsistente. Tais evidências são encontradas nos aportes lexicais correntes nos falares regionais brasileiros e no português do Brasil como um todo. Identificados seus étimos prováveis ou precisos, chegamos até suas línguas de origem e a seus respectivos falantes. No entanto, é preciso não perder de vista certos fatores de natureza extralinguística que contribuíram para assentar as bases necessárias para a instalação dessa matriz africana como parte do processo de configuração da nação brasileira.

Inicialmente, a densidade demográfica estimada em cinco a oito milhões de africanos introduzidos para substituir o trabalho escravo ameríndio, o que originou um contingente populacional de 75% de negros e mestiços em relação ao número de portugueses e outros europeus, conforme o censo oficial de 1823, um ano após a independência do Brasil. Essa vantagem, em termos de superioridade demográfica no confronto das relações de trabalho e

na convivência diária, contribuiria para o desaparecimento, já no século XVIII, de uma língua franca de base indígena, dando lugar à emergência de dialetos afro-brasileiros nas senzalas, plantações e minas, nos quilombos e, mais tardiamente, em núcleos urbanos da costa, até o seu estabelecimento em línguas litúrgicas e, dispondo de um vocabulário menor, sob a forma de falares especiais de comunidades negras isoladas. (...)

A análise dos dados (...) nos levou a concluir que, historicamente, por parte das línguas africanas, as do grupo banto foram as mais importantes no processo de configuração do perfil do português brasileiro, devido à antiguidade e superioridade numérica de seus falantes e à grandeza da dimensão alcançada pela sua distribuição humana no Brasil colonial. Assim sendo, na medida em que a profundidade sincrônica revela uma antiguidade diacrônica, constatamos que os aportes bantos estão associados ao regime da escravidão, aqueles chamados “empréstimos arcaicos” por Jacques Raimundo, alguns já obsoletos (cf. *mucama*), enquanto a maioria deles está completamente integrada ao sistema linguístico do português, formando derivados portugueses a partir de uma mesma raiz banto (cf. *molambo*, *esmolambar*, *esmolamento* etc.). Em alguns casos, a palavra banta chega a substituir completamente o seu equivalente em português, sem que o locutor brasileiro, em geral, seja capaz de discernir se aquela palavra é africana, ameríndia ou portuguesa.

Dentre os exemplos, merece destaque a palavra *caçula*, por ser a única conhecida de todos os brasileiros com o sentido de “filho mais jovem” e cuja origem africana é completamente ignorada pela grande maioria, um fato que vem corroborar a tese da importância do desempenho sociolinguístico da mulher negra servindo de “mãe-preta” na intimidade da família colonial, a começar da criança, e na condição, também, de escrava ladina. Para marcar mais ainda a extensão e profundidade do dimensionamento psicossocial da sua atuação, o *caçula* continua sendo visto, através de uma expressão muito popular no Brasil, como o “dengo da família”, ou seja, aquele sempre mimado e cheio de vontades, enfim, herdeiro do *dengo*, na voz africana de quem o criou.

Quanto a línguas da família kwa na região do Golfo de Benin, as do grupo ewe-fon foram registradas em Vila Rica (Ouro Preto), Minas Gerais, em um caderno redigido, no início do séc. XVIII, por Antônio da Costa Peixoto, mas só publicado em 1945, em Lisboa. Dos 831 termos que contém, identificamos 80% fon contra 20% mahi, gun, mina ou ewe, cujo conhecimento se revelou, portanto, anterior ao da língua iorubá no Brasil. Essa última, mais do que as outras, está concentrada nos aspectos religiosos da sua cultura e com pouca produtividade na formação de derivados portugueses, um tipo de dado que denuncia uma importação relativamente mais recente e coincide com a informação histórica quanto à introdução, em levadas numerosas e sucessivas, de seus falantes na cidade do Salvador, a partir da destruição do reino nagô de Queto, em fins do século XVIII, e do império iorubá de Oió, em 1830, até a extinção definitiva do tráfico transatlântico para o Brasil duas décadas depois. Segundo estatísticas aduaneiras levantadas por Viana Filho, no século XIX, a Bahia importou 850.000 oeste-africanos e 350.000 bantos. (...)

O que mais chamou nossa atenção foi constatar que, na maioria dos casos estudados, ocorria uma adaptação morfológica (morfemas de gênero e número) mais do que uma

evolução fonética das palavras importadas, diante das semelhanças casuais, mas notáveis, do sistema linguístico das línguas banto e kwa (...) com o sistema do português brasileiro. Entre elas, sete vogais orais, reconstituídas no protobanto e próprias do iorubá e do fon, que também conhecem as cinco vogais nasais, e, com exceção da nasal silábica para as línguas africanas, a estrutura silábica (CV), onde a vogal é sempre centro de sílaba, estabelecendo a fórmula (CV.CV) como representante da estrutura ideal, o que provavelmente possibilitou a continuidade do tipo prosódico de base vocálica do português arcaico na modalidade brasileira, afastando-o, portanto, da pronúncia atual, muito consonantal, do português europeu.

Não estaria, então, nessa proximidade relativa, e possivelmente em outras ainda encobertas, o elo perdido para fechar uma questão lógica quanto à avaliação da parte do influxo de línguas africanas no português do Brasil, se o Brasil possui a maior população de descendência negra concentrada fora do continente africano? Ainda o fato de não haver sucedido um crioulo brasileiro adquirido como segunda língua ou como língua nacional, semelhante às que emergiram em outras ex-colônias americanas, já que o português foi imposto, de qualquer maneira, como um falar estrangeiro a uma população majoritariamente de falantes africanos por três séculos consecutivos? Só um milagre seria capaz de responder a essa indagação, ainda mais quando é admissível que falantes de qualquer língua, por mais resistentes a mudar hábitos articulatórios da sua língua materna, tendem a acomodá-los ao sistema fonológico da nova língua adquirida, como se deve ter passado no Brasil em relação ao falante africano ante o português.

Como milagres desse tipo não acontecem, chegamos necessariamente a uma hipótese compatível com as circunstâncias extralinguísticas que foram favoráveis a este processo: o português do Brasil, naquilo em que se afastou, na fonologia, do português de Portugal é, *a priori*, o resultado de um compromisso entre duas forças dinamicamente opostas e complementares, ou seja, por um lado, uma imantação dos sistemas fônicos africanos em direção ao sistema do português e, em sentido inverso, um movimento do português em direção aos sistemas fônicos africanos, sobre uma matriz indígena preexistente e mais localizada no Brasil. Consequentemente, o português de Portugal, arcaico e regional, foi ele próprio, de certa forma, mais ou menos africanizado pelo fato de uma longa convivência. A complacência ou resistência face a essas influências mútuas é uma questão de ordem sociocultural e os graus de mestiçagem linguística correspondem, mas não de maneira absoluta, aos graus de mestiçagem biológica que se processam no país.

Guia de Leitura texto 9: As línguas da África

- (1) Quais as duas principais razões para a diversidade étnica africana?
- (2) Como se distribuíam, geograficamente, os grupos étnicos africanos por volta de 1400?
- (3) Qual a língua falada em Madagascar e qual a sua filiação linguística?
- (4) Quais as quatro grandes famílias linguísticas propostas por Joseph Greenberg para classificar as línguas africanas?
- (5) Qual a única língua nilo-saariana com tradição escrita?
- (6) Qual das quatro famílias é a mais expandida territorialmente ao longo do continente africano? Qual seu grupo mais importante em termos de extensão geográfica e número de falantes?
- (7) Os grupos nilo-saariano e coissã não se apresentam em áreas contínuas no mapa. O que tal fato pode sugerir com relação à evolução histórica e geográfica dos povos que falam essas línguas?
- (8) Como os estudos de Greenberg modificaram a antiga percepção sobre a origem das línguas semíticas?
- (9) Qual a consequência da expansão congo-cordofiana, segundo Greenberg, para os pigmeus?
- (10) Qual a característica sonora mais marcante das línguas coissãs?
- (11) Explique porque se acredita que a expansão banta não deva ser um acontecimento recente – sem, porém, ser um fato ancestral ou antiquíssimo.
- (12) A qual grupo pertence a maior parte das línguas dos negros trazidos para o Brasil?
- (13) Cite cinco palavras que hoje compõem o léxico do português e que advieram do quimbundo ou quicongo.
- (14) A completa integração ao português de termos vindos do quimbundo ou quicongo pode ser atestada com base em qual fato linguístico?
- (15) Como se divide o sistema lexical das línguas bantas? Em que se difere da diferenciação das classes de palavra do português?
- (16) Explique e exemplifique o funcionamento do genitivo em quimbundo.
- (17) Dê exemplos de contribuições lexicais das línguas kwa e iorubá para o português.
- (18) Atualmente quais os contextos em que o iorubá ainda é falado na Bahia?
- (19) Explique as especificidades morfológicas do iorubá quanto à categoria de gênero.
- (20) Explique as especificidades morfológicas do iorubá quanto à categoria de número.
- (21) Por que a influência das línguas africanas no Brasil permanece objeto de intenso debate entre os estudiosos?
- (22) É possível rastrear o influxo das línguas africanas no português brasileiro em um plano linguístico além do lexical?
- (23) Identifique e explique três tendências do português brasileiro que podem ser creditadas ao contato com as línguas africanas.

TEXTO 10

LÍNGUAS INDÍGENAS BRASILEIRAS

Quesler Fagundes Camargos

Selmo Azevedo Apontes

De acordo com o Instituto Brasileiro de Geografia e Estatística (2010), doravante IBGE, há no Brasil 817.963 índios autodeclarados, dos quais 315.180 vivem em área urbana e 502.783 em área rural. O estado com a maior quantidade de indígenas é o Amazonas com 168.680 indivíduos, enquanto que o estado com a menor quantidade de índios é o Rio Grande do Norte, com 2.597 pessoas. O estado de Minas Gerais possui 31.112, o que o coloca como o 11º estado com a maior população indígena brasileira.

Apesar de o censo realizado pelo IBGE em 2010 mostrar que os povos indígenas brasileiros integram 305 etnias registradas, as quais falam 274 línguas, linguistas (cf. RODRIGUES, 1986, 2002; NOLL; DIETRICH, 2010, entre outros) e associações linguísticas brasileiras (cf. Instituto Socioambiental²⁸ e Museu Goeldi²⁹, entre outras) afirmam que, no Brasil, são faladas, na verdade, aproximadamente 180 línguas indígenas.

Na verdade, há uma grande controvérsia quanto à quantidade de línguas indígenas brasileiras pelo fato de haver certa dificuldade em delimitar um dialeto e uma língua. Deve-se ressaltar ainda que esses números não incluem as línguas dos índios isolados, os quais, por estarem sem contato com a sociedade, não puderam ainda ser identificadas.

Para mais detalhes demográficos, convidamos o leitor a acessar o site <http://indigenas.ibge.gov.br>, onde o IBGE fornece mais informações sobre a distribuição da população autodeclarada indígena no território brasileiro, com base nos resultados censitários. Esta página foi criada pelo IBGE, em parceria com a FUNAI, no dia 19 de abril de 2012 em comemoração ao Dia do Índio.

Apesar de hoje haver apenas 180 línguas indígenas vivas no Brasil, estima-se, conforme Rodrigues (2005), que, na época do descobrimento do Brasil, existiam cerca de 1.200 línguas indígenas diferentes. Durante os últimos 500 anos, mais de 1 mil desses idiomas se perderam por diversos motivos, dos quais se pode citar: morte de índios e populações inteiras em decorrência de epidemias, extermínio, escravização e aculturação forçada.

Na atualidade, os povos indígenas, os quais sobreviveram a esse longo massacre, ainda são alvos de perseguições e sofrem constantes represálias principalmente no tocante à questão da terra. Rodrigues (2005, p. 36), por exemplo, afirma que:

²⁸ <http://pib.socioambiental.org/pt/c/no-brasil-atual/linguas/introducao>

²⁹ http://saturno.museu-goeldi.br/lingmpep/portal/?page_id=205

a redução de 1.200 para 180 línguas indígenas nos últimos 500 anos foi o efeito de um processo colonizador extremamente violento e continuado, o qual ainda perdura, não tendo sido interrompido nem com a independência política do país no início do século XIX, nem com a instauração do regime republicano no final desse mesmo século, nem ainda com a promulgação da “Constituição Cidadã” de 1988. Embora esta tenha sido a primeira carta magna a reconhecer direitos fundamentais dos povos indígenas, inclusive direitos linguísticos, as relações entre a sociedade majoritária e as minorias indígenas pouco mudou.

O primeiro contato feito com a população indígena no território brasileiro com a finalidade de estudar suas línguas foi realizado pelos missionários jesuítas, quando da colonização iniciada em abril de 1500. Nesta ocasião, o contato inicial foi feito com as tribos que habitavam o litoral: os Tupinambás e/ou os Tamoios, principalmente. Pelo fato de as línguas faladas no litoral serem muito semelhantes entre si, elas passaram a ser consideradas como o padrão das línguas indígenas faladas no território. A consequência desse fato histórico é que o Tupinambá foi uma das poucas línguas estudadas nos primeiros trezentos anos de colonização. A primeira gramática descritiva de uma língua indígena em solo brasileiro foi feita pelo Padre José de Anchieta, a qual foi publicada em 1595 com o título *Arte de Gramática da Língua mais usada na costa do Brasil*.

1 Famílias e troncos linguísticos

Assim como as demais línguas do mundo, as línguas indígenas faladas no Brasil, também conhecidas como línguas ameríndias brasileiras, por apresentarem semelhanças nas suas origens, se tornam parte de grupos linguísticos que são as famílias linguísticas. Estas últimas, por sua vez, fazem parte de grupos ainda maiores, os quais são denominados como troncos linguísticos. Nos troncos se agrupam as línguas cuja origem comum vem de milhares de anos, sendo as semelhanças entre elas muito sutis. Já nas famílias, as semelhanças são maiores, uma vez que as separações ocorreram há menos tempo.

Os troncos com maior número de línguas no Brasil são o Tupí e o Jê. Além desses dois troncos, há ainda aproximadamente 20 famílias linguísticas, as quais, por não possuírem taxas/quantidades suficientes de semelhanças, não puderam ser identificadas como relacionadas aos troncos Tupí e Jê e também não são agrupadas em outro tronco linguístico. Além dessas, existem também outras línguas que não puderam ser classificadas dentro de nenhuma família, permanecendo, assim, dentro da categoria de línguas isoladas. Veja o quadro abaixo, o qual apresenta a quantidade aproximada de línguas pertencentes a cada um desses troncos e famílias:

QUADRO 1 – Troncos e Famílias de línguas faladas no Brasil³⁰

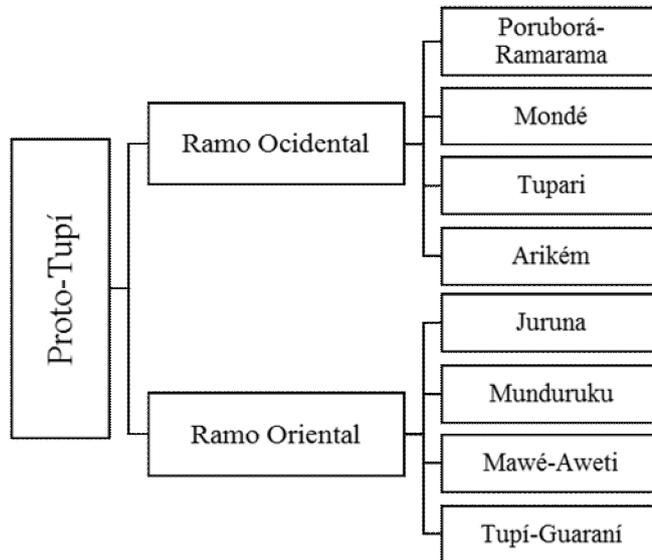
TRONCO LINGUÍSTICO	FAMÍLIA LINGUÍSTICA	
Tupí (10 famílias, 40 línguas)	Aikaná (1 língua)	Máku (1 língua)
	Arawá (8 línguas)	Makú (6 línguas)
Jê (16 famílias, 20 línguas)	Arúak (16 línguas)	Mura (2 línguas)
	Guaikuru (1 língua)	Nambikwára (3 línguas)
	Iranxe (1 língua)	Pano (12 línguas)
	Jabutí (2 línguas)	Trumái (1 língua)
	Kanoê (1 língua)	Tikúna (1 língua)
	Karib (20 línguas)	Tukano (11 línguas)
	Katukína (4 línguas)	Txapakúra (4 línguas)
	Koazá/kwazá (1 língua)	Yanomami (4 línguas)

2 Tronco linguístico Tupí

De acordo com Rodrigues (1958, 1985, 1986), Rodrigues e Cabral (2002) e Dietrich (2010), o tronco Tupí é constituído pelas seguintes famílias linguísticas geneticamente relacionadas: Tupí-Guaraní, Mondé, Arikém, Tuparí, Juruna, Mundurukú, Mawé-Aweti e Poruborá-Ramarama, conforme o seguinte organograma:

³⁰ A quantidade de línguas indicadas no quadro 1 é aproximada, visto que a definição de língua e de dialeto difere entre os autores.

ORGANOGRAMA 1 – Tronco Tupí³¹



Fonte: RODRIGUES, 1985, 1986; RODRIGUES; CABRAL, 2002; DIETRICH, 2010

A fim de determinar a filiação genética das línguas humanas, geralmente se considera tanto as estruturas das línguas quanto o seu léxico. Dentre essas propriedades, Rodrigues (1985, 1986) seleciona algumas como diagnósticos não só para incluir algumas línguas na família Tupí-Guaraní, mas também excluir as línguas geneticamente aparentadas, só que em nível mais remoto. Os diagnósticos escolhidos por Rodrigues (1985) pertencem exclusivamente ao campo lexical. O autor justifica essa abordagem com base no fato de que os estudos de línguas indígenas brasileira ainda são muito incipientes no campo da linguística.

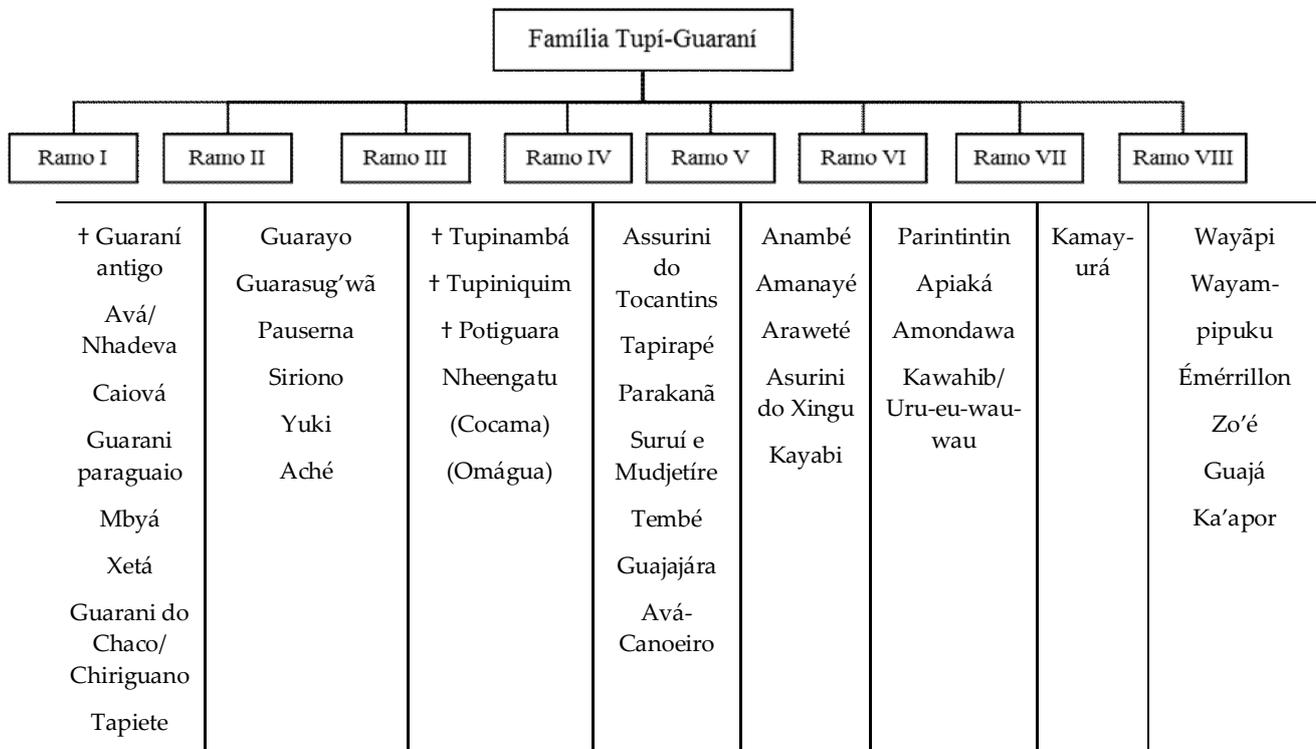
Rodrigues (1985, p. 48) ressalta que os subconjuntos (ramos) da família Tupí-Guaraní constituem

não propriamente uma classificação interna da família Tupí-Guaraní, mas antes um ensaio de discriminação de seções dessa família caracterizadas pelo compartilhamento de algumas propriedades linguísticas, as quais podem servir para diagnosticar o desmembramento de todo o conjunto de línguas Tupí-Guaraní visto como resultante histórico de uma proto-língua pré-histórica.

O resultado do estudo acima é apresentado esquematicamente no seguinte organograma:

³¹ As línguas da família Tupí-Guaraní são faladas em várias regiões brasileiras, além de outros países da América do Sul. Todas as outras famílias desse tronco estão situadas no Brasil, especificamente no sul do Rio Amazonas.

ORGANOGRAMA 2 – Subgrupos da família Tupí-Guaraní³²



Fonte: RODRIGUES, 1985, 1986; RODRIGUES; CABRAL, 2002; DIETRICH, 2010.

A família Tupí-Guaraní se destaca das demais famílias pela notável extensão territorial sobre a qual estão distribuídas suas línguas. Para se ter uma ideia, falam-se línguas dessa famílias no Amapá, no Amazonas, no Espírito Santo, em Goiás, no Maranhão, no Mato Grosso, no Mato Grosso do Sul, no Pará, no Paraná, no Rio de Janeiro, no Rio Grande do Sul, em Santa Catarina e em São Paulo, assim como fora do Brasil: na Guiana Francesa, na Venezuela, na Colômbia, no Peru, na Bolívia, no Paraguai e na Argentina. Para ver mais detalhes acerca das migrações dos povos Tupí, o que ocasionou essa significativa dispersão, convido o leitor a consultar o trabalho de Fausto (2005).

3 Tronco linguístico Jê

O termo Jê, conforme Ribeiro (2006), é relativo à família de mesmo nome e provavelmente é derivado do morfema coletivo [ʒe] das línguas dessa família. A terminologia Macro-Jê, por sua vez, de acordo com Rodrigues (1999), foi proposta por Mason (1950) a fim de designar um conjunto de línguas que, a princípio, estavam relacionadas à

³² O símbolo (†) marca uma língua extinta.

família linguística Jê. Embora o tronco linguístico Jê seja ainda uma hipótese de trabalho, como afirma Rodrigues (1999), ele é assumido por muitos autores. Dentre eles, destacamos Rodrigues (1986, 1999) e Ribeiro (2006).

De acordo com Rodrigues (1999), a grande parte das línguas que provavelmente compõe o tronco linguístico Jê já está extinta e as línguas que se mantêm vivas ainda são pouco documentadas. Devido a estes dois fatores, as relações internas desse tronco precisam ser significativamente investigadas, uma vez que ainda são hipotéticas. Para se ter uma ideia, não há um consenso na literatura sobre quantas famílias constitui o tronco Jê: a quantidade varia entre quinze e dezesseis.

De acordo com Greenberg (1987) e Ribeiro (2005), pertencem ao tronco Jê as famílias Jê, Boróro, Chiquitano, Guató, Jabuti, Kamakã, Karajá, Krenak, Maxakalí, Ofayé, Otí, Purí (Coroado), Rikbaktsá e Yatê. Greenberg (1987) e Rodrigues (1999) ainda defendem a inclusão da família Karirí. Veja o quadro abaixo que apresenta as 15 famílias desse tronco:

QUADRO 2 – Famílias e línguas do tronco Jê

	FAMÍLIAS	LÍNGUAS
1	Boróro	Boróro, †Umutína, †Otúke
2	Chiquitano	Chiquitano (Besiro)
3	Guató	Guató
4	Jabuti	Djeoromitxi (Jabuti), Arikapú
5	Jê	†Jeikó Jê setentrional: Panará, Suyá, Kayapó, Timbira (Parkatêjê, Pykobjê, etc), Apinajé Jê central: Xavante, Xerente, †Acroá-Mirim, †Xacriabá Jê austral: Kaingáng, Xoklém, †Ingain
6	Kamakã	†Kamakã, †Mongóyó, †Meniém, †Kotoxó, †Massakará
7	Karajá	Karajá (Karajá austral, Karajá setentrional, Javaé e Xambioá)
8	Karirí	Karirí, †Kipeá, Dzubukuá, †Pedra Branca, †Sabuyá
9	Krenak	Krenak (Botocudo, Borúm)
10	Maxakalí	†Pataxó, †Kapoxó, †Monoxó, †Makoní, †Malalí, Maxakalí
11	Ofayé	Ofayé
12	Otí ³³	†Otí (Eo-Xavante)
13	Purí (Coroado)	†Coroado, †Purí, Koropó
14	Rikbaktsá	Rikbaktsá
15	Yatê	Yatê

Fonte: GREENBERG, 1987; RIBEIRO, 2005; RODRIGUES, 1999.

³³ A família Otí é proposta de Greenberg (1987).

De acordo com Ribeiro (2006), as línguas pertencentes ao tronco Jê apresentam, em comparação com outras línguas indígenas brasileiras e de modo geral, as seguintes propriedades:

- (a) Inventário maior de vogais em relação ao inventário das consoantes.
- (b) Padrões silábicos mais simples do que àqueles de outras línguas indígenas.
- (c) Acento predizível.
- (d) Morfologia flexional relativamente simples.
- (e) O verbo geralmente se localiza no final da sentença.
- (f) Ocorrência de posposições em vez de preposições.
- (g) Os adjetivos são expressos por nomes ou por verbos descritivos.

Apresentamos, no quadro abaixo, o qual foi retirado de Rodrigues (2002, p. 50), algumas evidências do parentesco que une as línguas do tronco Jê como um conjunto. Mais precisamente, mostraremos semelhanças e padrões que demonstram que estas línguas possuem historicamente uma origem comum.

QUADRO 3 – Comparação de algumas línguas do tronco Jê

LÍNGUAS ³⁴ DO TRONCO JÊ	PÉ	UM	BRAÇO	FLECHA	MEL	FÍGADO	CINZA	MARIDO
Apinayé (5)	<i>par</i>	<i>pitxi</i>	<i>pa</i>	--	<i>mèñ</i>	<i>ma</i>	<i>mrò</i>	<i>mien</i>
Xavánte (5)	<i>paara</i>	--	<i>pano</i>	--	<i>pĩ</i>	<i>pa</i>	--	--
Kaingáng (5)	<i>pēn</i>	<i>pir</i>	<i>pē</i>	(<i>puñ</i>)	<i>mỹng</i>	<i>ta-mē</i>	<i>mrēi</i>	<i>mèn</i>
Maxakalí (10)	<i>pata</i>	<i>pytxèt</i>	--	<i>pói</i>	<i>pang</i>	--	<i>pytok</i>	<i>pen</i>
Kamakã (6)	<i>wade</i>	<i>weto</i>	--	<i>wãi</i>	--	--	--	--
Purí (13)	<i>txapere</i>	<i>i-pàin</i>	--	<i>pun</i>	--	--	--	--
Botocudo (9)	<i>pò</i>	<i>putxik</i>	<i>pò</i>	--	<i>pàng</i>	<i>ku-pagn</i>	--	--
Yatê (15)	<i>fe, fet-</i>	<i>fathowa</i>	--	--	--	--	<i>felowa</i>	(<i>feto</i>)
Kipeá (8)	<i>by, byri</i>	<i>bihe</i>	<i>bo</i>	<i>buiku</i>	--	--	<i>bydi</i>	--
Karajá (7)	<i>waa</i>	--	--	<i>wyhy</i>	<i>bâdi</i>	<i>baa</i>	<i>bry-by</i>	--
Boróro (1)	<i>byre</i>	(<i>mito</i>)	--	(<i>boi-</i>)	--	--	--	(<i>imedo</i>)
Ofayé (11)	<i>fara</i>	--	<i>fê</i>	--	<i>fyk</i>	<i>fa</i>	--	--
Guató (3)	<i>bò</i>	--	<i>pò</i>	--	<i>pagwa</i>	<i>pè</i>	--	--
Rikbaktsá (14)	<i>pyry</i>	--	<i>txi-pa</i>	--	<i>mêk-</i>	--	--	<i>mari-hta</i>

³⁴ A numeração que acompanha o nome de cada língua do tronco Jê corresponde à família linguística indicada no quadro 2 da pág. 6.

Segundo Rodrigues (2002), a série para a palavra “pé” é praticamente a única com palavras cognatas e documentadas em todas as línguas apresentadas no quadro acima. São tidas como possivelmente cognatas porque sua constituição fonética permite supor que sejam todas derivadas por modificações de uma só palavra de uma língua ancestral. Provavelmente, o primeiro som dessa palavra ancestral é uma consoante labial, como /p/ ou /b/, o segundo som é uma vogal central, como /a/ ou /y/, e o terceiro som é uma consoante dental, como /r/, /d/ ou /t/. Note ainda que essas modificações de sons se repetem regularmente em outras séries: as consoantes iniciais de “pé” são as mesmas de “um”, de “braço” e de “flecha”. Além do mais, note que as palavras cognatas das línguas Apinayé, Xavánte e Kaingáng são foneticamente mais semelhantes entre si do que com as outras línguas do tronco Jê. Essa maior aproximação entre essas três línguas se justifica, uma vez que se tratam de línguas de uma mesma família linguística: a família Jê.

4 Família linguística Karíb

Conforme Rodrigues (2002), o nome Karíb (Caribe) é uma das designações pelas quais foi conhecido um povo indígena que habitou, nos séculos passados, grande parte da costa norte da América do Sul e as Pequenas Antilhas. Seu território se estendia desde o norte da foz do rio Amazonas, atravessando a Guiana Francesa, o Suriname e a Guiana, até chegar à Venezuela.

A maior parte das línguas dessa família faladas em território brasileiro está situada ao norte do rio Amazonas, no Amapá, no norte do Pará, em Roraima e no Amazonas. Há ainda línguas Karíb ao sul do rio Amazonas, situadas essencialmente ao longo do rio Xingu. O quadro abaixo, retirado de Rodrigues (2002, p. 58), apresenta uma pequena amostra que compara línguas do norte do rio Amazonas.

QUADRO 4 – Comparação de línguas Karíb faladas ao norte do rio Amazonas

	Galibí	Apalaí	Wayâna	Hixka-ryâna	Taulipáng
LUA	<i>nuno</i>	<i>nuno</i>	<i>nunuy</i>	<i>nuno</i>	<i>kapyi</i>
SOL	<i>wéiu</i>	<i>xixi</i>	<i>xixi</i>	<i>kamymy</i>	<i>wéi</i>
ÁGUA	<i>tuna</i>	<i>tuna</i>	<i>tuna</i>	<i>tuna</i>	<i>tuna, paru</i>
CHUVA	<i>konopo</i>	<i>konopo</i>	<i>kopo</i>	<i>tuna</i>	<i>kono'</i>
CÉU	<i>kapu</i>	<i>kapu</i>	<i>kapu</i>	<i>kahe</i>	<i>ka'</i>
PEDRA	<i>topu</i>	<i>topu</i>	<i>tepu</i>	<i>tohu</i>	<i>ty'</i>
FLECHA	<i>pyrywa</i>	<i>pyróu</i>	<i>pyréu</i>	<i>waiwy</i>	<i>pyrýu</i>
COBRA	<i>okóiu</i>	<i>âkóia</i>	<i>ykýia</i>	<i>okóie</i>	<i>ykýi</i>

PEIXE	<i>wuoto</i>	<i>kana</i>	<i>kaa</i>	<i>kana</i>	<i>moro'</i>
ONÇA	<i>kaituxi</i>	<i>kaikuxi</i>	<i>kaikui</i>	<i>kamara</i>	<i>kaikuse</i>

De acordo com Rodrigues (2002), as línguas que não foram incluídas no quadro acima estão associadas mais diretamente com uma ou com outra dessas cinco. O Makuxí e o Ingarikó, por exemplo, estão intimamente ligadas ao Taulipáng, com palavras semelhantes para “lua”, “sol” e “peixe” e com a perda da sílaba final *-pu* ou *-po*, substituída por uma oclusão glotal (’), nas palavras “chuva”, “céu” e “pedra”. Além do mais, observe que essas línguas são fortemente aparentadas, o que justifica incluí-las em uma mesma família linguística.

Agora, no seguinte quadro, retirado de Rodrigues (2002, p. 59), apresentamos as línguas Karíb do sul do rio Amazonas.

QUADRO 5 – Comparação de línguas Karíb faladas ao sul do rio Amazonas

	Galibí	Apalaí	Wayâna	Hixka-ryâna	Taulipáng
LUA	<i>nuno</i>	<i>nuno</i>	<i>nunuy</i>	<i>nuno</i>	<i>kapyi</i>
SOL	<i>wéiu</i>	<i>xixi</i>	<i>xixi</i>	<i>kamymy</i>	<i>wéi</i>
ÁGUA	<i>tuna</i>	<i>tuna</i>	<i>tuna</i>	<i>tuna</i>	<i>tuna, paru</i>
CHUVA	<i>konopo</i>	<i>konopo</i>	<i>kopo</i>	<i>tuna</i>	<i>kono'</i>
CÉU	<i>kapu</i>	<i>kapu</i>	<i>kapu</i>	<i>kahe</i>	<i>ka'</i>
PEDRA	<i>topu</i>	<i>topu</i>	<i>tepu</i>	<i>tohu</i>	<i>ty'</i>
FLECHA	<i>pyrywa</i>	<i>pyróu</i>	<i>pyréu</i>	<i>waiwy</i>	<i>pyryú</i>
COBRA	<i>okóiu</i>	<i>âkóia</i>	<i>ykýia</i>	<i>okóie</i>	<i>ykýi</i>
PEIXE	<i>wuoto</i>	<i>kana</i>	<i>kaa</i>	<i>kana</i>	<i>moro'</i>
ONÇA	<i>kaituxi</i>	<i>kaikuxi</i>	<i>kaikui</i>	<i>kamara</i>	<i>kaikuse</i>

Uma classificação interna da família Karíb, puramente linguística, depende de uma análise que considere o comportamento dos sons, da gramática e do vocabulário. De acordo com Rodrigues (2002), este estudo, o qual foi iniciado no século passado por Steinen (1886, 1892) e Adam (1893), ainda se mantém em um estado rudimentar, devido à precariedade da documentação existente. Para sermos mais exatos, somente nos últimos 30 anos houve um número crescente de trabalhos linguísticos sobre as línguas Karíb. As línguas que receberam certa atenção dos linguistas foram: o Waiwai (HAWKINS, 1998), o Makuxi (ABBOTT, 1991; MACDONNELL, 1994), o Apalaí (KOEHN; KOEHN, 1986), o Wayana (JACKSON, 1972; TAVARES, 2005), o Panare (MATTEI MULLER, 1994), o Ye'kwana (HALL, 1988), o Tiriyo

(MEIRA, 1999; CARLIN, 2004), o Bakairi (SOUZA, 1994), o Ikpeng (PACHÊCO, 2001) e o Arara (SOUZA, 1993), os quais são essencialmente dissertações e teses de pós-graduação.

As classificações da família Karíb mais recentemente publicadas foram os trabalhos de Girard (1971), Durbin (1977), Kaufman (1994) e Meira e Franchetto (2005). Para Rodrigues (2002), deve-se, no entanto, levar em consideração que estes autores apresentam vários pontos duvidosos e, em alguns casos, conflitantes, uma vez que não estão de acordo sobre o grau de parentesco entre algumas línguas. Por exemplo, a língua Waimiri-Atroari é considerada ora como língua isolada dentro da família ora faz parte do subgrupo junto com o Mapoyo e o Yawarana. A principal razão para esta disparidade é a falta de dados que permita uma conclusão definitiva. Dentre todas as propostas já mencionadas, apresentamos no quadro abaixo a classificação proposta por Meira e Franchetto (2005) por dois motivos: (i) porque revê as classificações anteriores e (ii) porque é a classificação mais recente. Contudo, alertamos o leitor para o fato de que ainda não é uma classificação definitiva, uma vez que há algumas línguas pouco conhecidas.

QUADRO 6 – Classificação das línguas Karíb

RAMO	GRUPO	LÍNGUA
Guianense	Taranoano	Tiriyó
		Akuriyó
		Karihona
	Parukotoano	Waiwai
		Hixkaryana
		Katxuyana
		Karinya (Galibi) Wayana Apalaí (?) † Palmella (?)
Venezuelano	Costeiro	† Chayma † Chumanagoto
	Pemonguiano	Pemong (Arekuna, ...) Kapong (Akawaio, ...) Makuxi
		Panare Ye'kwana (?)
		Mapoyo (?) Yawarana (?)
		† Tamanaku
Waimiriano		Waimiri-Atroari (?)
Yukpano		Yukpa (Motilón) Hapreia (Japreria)
Sul (ou Pekodiano)	Xiguano	Arara Ikpeng
		Bakairi

Kuikuroano		Kuikuro (Kalapalo, ...) † Pimenteira (?)
------------	--	---------------------------------------------

5 Família linguística Pano

Segundo Ferreira (2005), no Brasil, a preocupação com o estudo das línguas indígenas reuniu quatro grandes troncos e famílias de línguas no país: Tupi, Jê, Aruak e Karib. No entanto, pequenos grupos, como o das línguas Pano, foram esquecidos (CÂMARA, 1979). Atualmente, mais pesquisadores estão se interessando pelas línguas Pano, o que resulta em um aumento gradual de estudos dessas línguas.

Em se tratando da filiação, Paula (2004) faz uma descrição da filiação e o histórico da questão: as línguas da família linguística Pano são faladas por povos indígenas que estão distribuídos em três países: Bolívia, Brasil e Peru. O nome da família foi dado por Grasserie (1890), quando, ao estudar um grupo de seis línguas - Caripuna, Conibo, Culino, Maxuruna (Mayoruna), Pakagwara e Wariapano (ou Pano) - descobriu semelhanças entre elas atribuindo o nome da última a todo o agrupamento linguístico. Brinton (1891), por sua vez, amplia o número de línguas para dezoito, propondo assim uma revisão na classificação proposta por Grasserie (1890).

Somente em Rivet e Tastevin (1927) é que a família Pano foi subdividida em três sub-grupos geográficos: o grupo I correspondia às línguas faladas nos rios Amazonas e Ucayali; o grupo II correspondia às línguas do rio Inambary; e, por fim, o grupo III ocupava as margens dos rios Mamoré, Beni e Madre de Dios. Após Rivet e Tastevin (1927), surgiram outras classificações, as quais foram importantes para o campo da linguística. Podemos citar os trabalhos de Nimuendaju (1932), Loukotka (1939), Mason (1950), Rivet e Loukotka (1952).

A classificação proposta por Mason (1950), embora tenha apresentado alguns problemas em relação aos nomes de alguns grupos (cf. KESINGER, 1985), caracterizou-se por sumarizar todas as classificações anteriormente propostas. O autor dividiu as línguas em três grupos. A sistemática adotada possibilitou uma nova organização das línguas em Pano Central, Pano Sul-Occidental e Pano Sul-Oriental.

Já d'Ans (1973) propõe uma reclassificação das línguas Pano, desmontando a classificação tradicionalmente aceita de Pano Central, Pano Sul-Oriental e Pano Sul-Occidental feita por Mason (1950). Esta última subdivisão é considerada por d'Ans (1973) como inexistente, tendo em vista que Mason (1950) se baseou em dados incorretos. A nova classificação proposta estabelece uma divisão das línguas da família Pano em cinco blocos, conforme o quadro abaixo:

QUADRO 7 – Classificação das línguas Pano

PANO UCAYALINO	PANO PRÉ-ANDINO	PANO DAS CABECEIRAS	PANO BENIANO	PANO DO NORTE?
Ucayalino A: Shipibo Conibo Capanahua	Cashibo Catacaibo?	Isconahua Amahuaca Cashinahua	Chácobo Pacaguara?	Mayoruna
Ucayalino B: Panavarro Shetebo Wariapano		Pano-Purus: Yaminahua Sharanahua Marinahua? Chaninahua? Mastanahua? Yahuanahua		

Ainda na década de 1970, houve um avanço considerável na classificação das línguas Pano, tendo as pesquisas realizadas no Peru como referência. As contribuições mais destacadas foram as de Shell (1975) e Loos (1975). Shell (1975) apresentou a primeira reconstrução dos proto-fonemas Pano, ao mesmo tempo em que descreveu alguns aspectos morfológicos como o marcador de transitividade, enquanto Loos (1975) apresentou trabalhos relacionados à morfologia destas línguas.

Mais recentemente, Loos (1999) produziu uma subclassificação da família Pano, baseando-se em características morfológicas e fonológicas, assim como em itens lexicais cognatos entre as línguas. Segundo este autor, as línguas da família linguística Pano apresentam muitas semelhanças, o que representa expansão e divisões recentes. Valenzuela (2003) alia uma avaliação das diferentes propostas de subclassificação da família linguística às suas próprias observações e conclui que ao menos seis agrupamentos diferentes devem ser postulados.

Do lado brasileiro, a subdivisão feita por Lanes (2005), utilizando o método léxico estatístico, ou seja, uma classificação que se baseia em critérios linguísticos e não geográficos, agrupou as línguas do ponto de vista lexical.

Segundo Valenzuela (2003), algumas das características tipológicas comuns a todas as línguas Pano são as seguintes:

- (a) Ordem básica dos constituintes: SOV/SV.
- (b) Uso exclusivo de sufixos e posposições (algumas línguas apresentam prefixos relativos a partes do corpo).
- (c) Ausência de marcação pronominal no verbo ou auxiliar.

(d) Ausência de concordância adnominal de qualquer tipo.

(e) Tendência dos marcadores de caso ocorrerem apenas uma vez, no último elemento do sintagma nominal.

A partir do fato de já haver uma quantidade suficiente de trabalhos descritivos das línguas da família Pano, os próximos passos são: (i) classificação léxico-estatística, a fim de verificar se as divisões que tomam por base a geografia serão sustentadas por dados linguísticos; e (ii) uma comparação interlinguística tendo por base a sintaxe, a fim de verificar se o padrão das ramificações proposto com base na morfologia se sustenta.

6 Família linguística Txapakura

De acordo com Rodrigues (1986), no vale do Guaporé e nos afluentes da margem direita do rio Madeira, no oeste de Rondônia e no sul do Amazonas, estendia-se até há não muito tempo uma das famílias linguísticas menos conhecidas, a família Txapakura, a qual não se incluem em nenhum dos grandes agrupamentos genéticos dos Tupí, Jê, Karib e Aruak. A ela, filiam-se as línguas do Pakaanóva e dos Urupá, em Rondônia, a dos Torá, no Amazonas, e também a dos Moré, na Bolívia. Já nos séculos passados, segundo o autor, os Torá eram os representantes mais setentrionais da família. Rodrigues (1986) afirmava que ainda não existia nenhum estudo científico das línguas da família Txapakúra, embora soubesse da presença dos missionários (Missão Novas Tribos) que conhecem e estavam analisando a língua dos Pakaa Nova.

Para demonstrarmos o parentesco linguístico das línguas pertencentes à família Txapakura, combinamos no quadro abaixo uma lista com cinco vocábulos das línguas Torá, Urupáe, Chapakura, Pawumwa, Iten e Oro Waram, os quais foram retirados de Hanseman (1912), Montfort e Rivet (1913), Nimuendaju e Bentes (1924), Nimuendaju (1925) e Loukotka (1963):

QUADRO 8 – Comparação vocabular entre algumas línguas Txapakura

	Torá	Urupá	Chapakura	Pawumwa	Iten	Oro Waram
DENTE	<i>iat</i>	<i>leti/yeti-si</i>	<i>yati-či</i>	<i>i'tiči</i>	<i>yia</i>	<i>yat</i>
LÍNGUA	<i>kapiak</i>	<i>kapiaka</i>	<i>kapikače</i>	<i>kabī'katci</i>	<i>kapaya</i>	<i>kapijaxi</i>
PAI	<i>ité</i>	<i>été</i>	<i>tia-tia</i>		<i>ité</i>	<i>ate/te</i>
MÃE	<i>inia</i>	<i>ive</i>			<i>ina</i>	<i>ina'</i>
OLHO	<i>tok</i>	<i>tyke/tũke-si</i>	<i>tuku-či</i>	<i>tu'kichi</i>	<i>tukichi</i>	<i>tok/tokoxi</i>

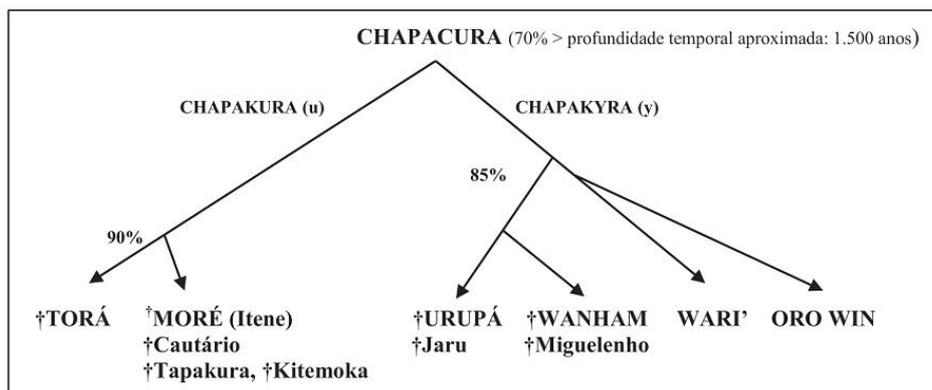
Segundo Meireles (1986), Rodrigues (1986) considerou o Txapakura como um dos principais grupos que vivem inteiramente dentro da Amazônia, mas o único que não foi objeto de estudos descritivos. Com os dados de que dispunha, como vimos anteriormente, considerou os povos que fazem parte da família Txapakura: os Pakaas-Novos, os Urupá e os

Moré. Claro é que a lista estava incompleta, porque os registros mais antigos das missões jesuíticas de Mojos (Bolívia) e dos relatos dos viajantes e naturalistas eram de difícil acesso.

Angenot-de-Lima (2001) apresenta 27 etnônimos para a família Txapakura. Porém, onze deles são apenas notas etnográficas: Jamari, Tapoaya, Kutiana, Matáwa, Urunamakam, Kujuna, Muré, Itoreauhip, Rokorona, Herisobokono, Kusikia-Manasi. Os outros 16 possuem registro vocabular: Torá, Urupá, Jarú, Oro Win, Kuyubi (kawtayo), Kumana, Uomo, Pawumwa, Abitana, Kabishi, Miguelenho, Moré, Chapakura (Tapakura), Kitemoka, Napeka, Oro Wari. Na lista dos provavelmente extintos, estão quase todos da lista, com ressalvas: Kuyubi (?), Oro Win, Moré, Oro Wari (com oito etnias).

Ramirez (2010) faz uma excelente análise dessa pequena família linguística e estabelece, por meio do método léxico-estatístico³⁵, uma proposta para o grupamento das línguas Txapakura, considerando que vários dos registros étnicos são etnônimos sinônimos como: Torá, Moré e Itene seriam codialetos. Os Abitana, Pawumwa, Miguelenho, Uomo e Kabishi, para o autor, são Wanyam. Os Cautário, Cujubim, Rokorono, Matawá, Kumaná, Itene e Herisobocono são Moré. Assim, o autor propôs a seguinte classificação.

ORGANOGRAMA 3 – Classificação das línguas da família Txapakura



Essa classificação se parece com a de Nimuendaju e Bentes (1922), que estipula o agrupamento ‘Chapakura-Wanham’: um representante à margem esquerda do Guaporé-Mamoré (lado boliviano) e outro à margem direita dos rios Guaporé-Mamoré-Madeira (do lado brasileiro). Ramirez (2010) situa os Cautário (Kawtayo, Kumaná) e os Torá, apesar de estarem à margem direita do Guaporé-Mamoré-Madeira, como mais próximo do Moré, pelo fato das listas de palavras serem mais idênticas e pelos registros dos viajantes corroborarem com a informação de que eram historicamente mais próximos.

Podemos ver claramente várias dessas línguas já extintas: Torá e Cautário (Brasil-BR) Tapakura, Kitemoka (Bolívia-BO), o Urupá e o Jarú (BR); ou em fase de extinção evidente como o caso do Moré (BO), Wanham, Miguelenho (BR). Outra língua que ainda se encontra

³⁵ E também baseado nos relatos de D’Orbigny (1843), Hanseman (1912), Créqui-Montfort e Rivet (1913), Nordenskiöld (1913-1914), Nimuendaju e Valle Bentes (1921), Metraux (1949), Loukotka (1963).

em uma situação bastante preocupante é o Oro Win (BR). Contando com aproximadamente 2.721 falantes (cf. FUNASA, 2006, vide ISA³⁶), está o grupo Wari', que mesmo não se situando entre as línguas em perigo (*endangered languages*), encontra-se em franca desvantagem com a língua oficial, em situação de vulnerabilidade.

O Grupo Wari' ou Pacaa Nova

A primeira referência desse grande grupo Pakaa Nova é da equipe de Rondon e Barbosa de Faria ([1927] 1948), os quais os situaram no trecho entre as cachoeiras do Ribeirão e Lage (no município de Nova Mamoré, RO). Depois, foi constatado que esse grupo se estendia até os rios próximos à cidade de Guajará-Mirim, rios Pacaás Novos e rio Sotério, em Rondônia. Os autores apresentaram uma lista de vocabulário, a partir da qual puderam situar o grupo como parte da família Txapakura.

Ryden (1954), ao estudar o grupo Moré (antigo Itenes), fez um resumo histórico dos relatos referentes aos povos que fazem parte dessa família Txapakura e confirma que os Pakaa Nova têm similaridades com a língua Moré. Depois foi estipulado que esse grande conjunto era, na verdade, composto de pequenos grupos com denominações próprias.

Não há consenso na denominação dos Wari', uma vez que este vocábulo significa "pessoa, gente, índio", e na denominação dos Oro Wari, que significa "nós indígenas" (*oro* = coletivizador). Para Kern e Everett ([1997] 2006), *wari'* significa 'nós' (pronome enfático de 1ª pessoa inclusiva). Também não há consenso em denominar o grande grupo de Pakaa Nova, pois é nome do primeiro rio acima da cidade de Guajará-Mirim.

Esse grupo, composto de um intrincado número de povos que flutuam em um estatuto entre língua e variante dialetal, é subdividido em duas partes, as quais são delimitadas por fronteira geográfica (tendo como barreira a Serra Pacaás Novos, continuação da Serra dos Parecis). Veja a subdivisão interna no seguinte quadro:

QUADRO 9 – Classificação do Grupo Wari'

WARI'	
Norte	Sul
Oro Mon	Oro Nao'
Oro Waram	Oro Eo
Oro Waram Xiyein	Oro At
Kao Oro Waye	Oro Yowin

Sabe-se que a classificação linguística que toma por base a proximidade geográfica levou a muitos equívocos na classificação e agrupamento de línguas, como pode ser observado em d'Ans (1973). A classificação acima situa o grande grupo no modo como as línguas foram identificadas à época do contato: décadas de 50 e 60 do século passado. Nesse período, viviam cada um com seu grupo. Após o período de contato, vários grupos passaram

³⁶ <http://pib.socioambiental.org/pt/c/quadro-geral>.

a viver juntos nas Terras Indígenas: Ribeirão, Igarapé Lage, Pacaás Novas, Rio Negro Ocaia e Sagarana. Vejamos um exemplo comparativo da estrutura silábica para a palavra *jacaré*:

QUADRO 10 – Estrutura silábica comparativa da palavra *jacaré*

JACARÉ	ESTRUTURA SILÁBICA	LÍNGUAS
<i>ahop</i>	CV.CVC	Moré
<i>ahop</i>	CV.CVC	Wa am/Miguelenho
<i>ahop</i>	V.CVC	Oro Nao'
<i>hop</i>	CVC	Oro waram
<i>βop</i>	CVC	Oro Win

Percebe-se que há um estágio gradativo de modificação do lexema, indo do Moré ao Oro Win. Parece que o Oro Nao' (grupo sul) e o Oro Waram (grupo norte) apresentam-se como um estágio intermediário entre o Moré e o Oro Win. Fato é que só será possível confirmar com descrições do grupo todo, para além dos dados de comparação de item vocabular.

7 Considerações finais

As línguas indígenas que ainda estão vivas no território brasileiro possuem uma grande diversidade linguística: desde a organização dos sistemas de sons até as estruturas gramaticais possíveis. Somente a partir da década de 80, houve um desenvolvimento significativo nos estudos da linguística indígena. Mesmo assim, atualmente, poucas instituições federais reservam um grupo de profissionais da área da linguística que seja dedicado a essas línguas. Entre elas, podemos citar: a Universidade de Brasília, a Universidade de Campinas, a Universidade de São Paulo, a Universidade Federal de Rio de Janeiro (Museu Nacional), o Museu Emílio Goeldi, entre algumas outras instituições que estão localizadas principalmente na região norte do Brasil. No entanto, apesar desses esforços, vale ressaltar que até agora poucas línguas indígenas foram estudadas com profundidade. Por essa razão, o conhecimento desses idiomas é constantemente atualizado. Quando ocorrem novas descrições, novas descobertas surgem, o que permite reformulações de teorias linguísticas.

Decidimos terminar esse texto citando trechos de Rodrigues (1966, p. 4-5),

as línguas indígenas constituem um dos pontos para os quais os linguistas brasileiros deverão voltar a sua atenção. Tem-se aí, sem dúvida, a maior tarefa da linguística no Brasil. [...] Cada nova língua que se investiga traz novas contribuições à linguística; cada nova língua é uma outra manifestação de como se realiza a linguagem humana; [...] cada nova estrutura linguística que se descobre pode levar-nos a alterar conceitos

antes firmados e pode abrir-nos horizontes novos para a visualização geral do fenômeno da linguagem humana. [...] Desde que se tenham algumas descrições de línguas, aparecerão espíritos curiosos bastante para dedicar-se a comparar essas descrições e daí tirar conclusões, classificando as línguas como relacionadas umas com as outras ou como pertencentes a tipos semelhantes num ou noutro particular, e para fazer deduções de ordem mais profunda, no âmbito da linguística geral e no campo das ciências antropológicas. [...] Se é lícito falar em responsabilidade de uma comunidade com respeito à investigação científica na região em que vive essa comunidade, então os linguistas brasileiros têm aí uma responsabilidade enorme, que é não deixar que se percam para sempre cento e tantos documentos sobre a linguagem humana.

Observe que aproximadamente 50 anos depois de publicado, a tarefa apontada por Rodrigues (1966) ainda está por fazer. De fato, das prováveis 180 línguas faladas atualmente no Brasil, há algumas que receberam pouca atenção de linguistas, o que resultou em alguns textos publicados, e muitas que ainda não possuem um estudo exaustivo, geralmente publicado por meio de gramáticas descritivas. Para se ter uma ideia do atual estado de descrição dessas línguas, citamos o trabalho de Moore (2007b), segundo o qual, de todas as línguas indígenas brasileiras, apenas 9% possuem uma descrição completa (i.e. descrição da gramática, coletânea de textos, dicionário); 23% apresentam uma descrição avançada (i.e. tese de doutorado ou muitos artigos); 34% possuem uma descrição incipiente (i.e. dissertação de mestrado ou alguns artigos); e 29% dessas línguas não possuem trabalhos com alguma importância científica. Além disso, segundo Moore (2007a), 23% das línguas brasileiras estão ameaçadas de extinção em curto prazo, por causa de seus números reduzidos de falantes e de baixa transmissão à nova geração.

A situação se apresenta de forma mais agravante quando examinamos os dados estatísticos concernentes à quantidade de índios por população indígena e à quantidade de falantes. Vale ressaltar que tais línguas sofrem uma forte pressão por parte da sociedade envolvente. Moore e Gabas (2006, p. 436) citam um estudo inicial feito por Morre (2000) indicando que:

no Brasil 24% das línguas são faladas por grupos com população de 50 pessoas ou menos; 8% por grupos com população de 51-100 pessoas; 25% por grupos com 101-250 pessoas; 18% por grupos na faixa de 201-500 pessoas; 10% por grupos com uma população estimada entre 501-1000 pessoas. Somente 15% do total de línguas brasileiras são faladas por grupos superiores a mil pessoas.

Agora, não se sabe a situação real dessas línguas em relação ao grau de transmissão. Moore e Gabas (2006, p. 436) exemplificam a situação de Rondônia, tendo por base o estudo inicial feito por Moore e Storto (1992), os quais estimaram que das 25 línguas do estado:

10% não estão mais em uso (como o caso do Puruborá); 30% têm um número baixo de falantes, e os jovens estão deixando de usar a língua; 25% ou têm número baixo de falantes ou faltam falantes jovens (mas não ambos os fatores concomitantemente); 35% têm falantes numerosos, incluindo jovens.

A partir dos dados acima, percebe-se que 65% das línguas do estado estão em situação crítica. Esse fato indica que podemos perder a contribuição da experiência desses povos com a linguagem.

Como convite, deixamos para o leitor a tarefa de investigar quais outras línguas indígenas brasileiras estão em processo acelerado de extinção (i.e. na iminência de desaparecer). Para isso, acesse o Atlas das Línguas do Mundo em Perigo da UNESCO, o qual pode ser acessado pelo site <http://www.unesco.org/culture/languages-atlas/index.php>. No site, encontrará informações mais específicas sobre as línguas e sua situação. Ademais, convidamos também o leitor a conhecer mais sobre os povos indígenas brasileiros no endereço eletrônico <http://pib.socioambiental.org/pt/c/no-brasil-atual/linguas/introducao>.

Guia de Leitura Texto 10: Línguas Indígenas Brasileiras

- (1) Segundo Rodrigues, qual a estimativa aproximada do número de línguas indígenas no território brasileiro na época do descobrimento? Qual a estimativa de línguas indígenas na atualidade?
- (2) Quais as primeiras línguas indígenas estudadas e descritas no Brasil?
- (3) Quais os troncos com o maior número de línguas?
- (4) Quais critérios são empregados, normalmente, para estabelecer-se a filiação genética das línguas humanas?
- (5) Cite algumas regiões ou zonas territoriais em que línguas da família tupi-guarani são faladas.
- (6) Enumere ao menos quatro características das línguas pertencentes ao tronco jê.
- (7) Qual a região ou porção territorial foi ocupada pelos povos Karíb nos séculos passados?
- (8) De que depende uma análise estrita e puramente lingüística das línguas da família Karíb? Em que estágio encontra-se esta análise atualmente?
- (9) Por quais países estão distribuídas as línguas da família Pano?
- (10) Qual a subdivisão proposta por Rivet e Tastevin para as línguas desta família?
- (11) Qual a proposta de subdivisão de d'Ans?
- (12) Cite ao menos quatro características tipológicas das línguas Pano.
- (13) Em que região são faladas as línguas da família Txapakura.
- (14) Qual a primeira referência do grupo Pakaa Nova?
- (15) Explique, sinteticamente, a situação atual do estudo das línguas indígenas brasileiras como um todo.

BIBLIOGRAFIA

ABBI, Anvita. *Vanishing Voices of the Great Andamaneses* (VOGA). Disponível em: <<http://andamanese.net>>. Acesso em: 18 jul. 2011.

ABBOTT, Miriam. Macushi. In: DERBYSHIRE, Desmond. C.; PULLUM, Geoffrey. K. (Eds.). *Handbook of Amazonian languages*. Berlin: Mouton de Gruyter, 1991. v. 3. p. 23-160.

ADAM, Lucien. *Matériaux pour servir à l'établissement d'une grammaire comparée des dialectes de la famille caribe*. Paris: Librairie-Éditeur J. Maisonneuve, 1893.

AJAYI, Tayo Julius. *Empréstimo e variação interlinguística: o iorubá em contato com o português no Brasil*. 2002. 424 f. Tese (Doutorado em Estudos linguísticos). Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2002. N° chamada (Letras): 418 A312e 2002 T

ALKMIN, Tania; PETTER, Margarida. Palavras de África no Brasil de ontem e de hoje. In: FIORIN, José Luiz; PETTER, Margarida (Orgs.). *África no Brasil*. São Paulo: Contexto, 2008. p. 145-177. N° de Chamada: 469.798 A258 2008 (FAFICH)

ALVES, António Bárbolo. Para quei sirbe l mirandês. Disponível em: <<http://mirandes.blogs.sapo.pt/2007/04>>. Acesso em: 18 jul. 2011.

AMBROSINI, Ricardo. *Introduzione alla glotologia indo-europea*. Pisa: ETS, 1998.

www.ancientscripts.com/. Acesso em: 23 abr. 2010.

ANCHIETA, Pe. José de. *Arte de Gramática da lingoa mais usada na costa do Brasil*. Ed. Fac-similar à de 1595. São Paulo: Loyola, 1990.

ANGENOT DE LIMA, Geralda. *Description phonologique, grammaticale et lexicale du moré, langue amazonienne de Bolivie et du Brésil*. Porto Velho: EDUFRO, 2002.

APONTES, Selmo A.; CAMARGOS, Quesler F. *Processos de Causativização em Oro Waram e suas consequências para a codificação dos argumentos nucleares*. Apresentação no IV CIELLA, Pará. 2013 (no prelo)

APONTES, Selmo A.; KEMPF, Chaterine Barbara. A problemática do dito “embasamento linguístico” no estabelecimento de uma ortografia: exemplificação Wari’. *Atas do I Encontro Internacional do GTLI da ANPOLL*, 2002. p. 343-352.

APONTES, Selmo A. *Descrição do Oro Waram, Variante Wari’ Norte (Pakka Nova, Txapakura)*. Artigo apresentado no Congresso da ABRALIN, Natal (RN). 2013a. (no prelo).

APONTES, Selmo A. *Processos fonológicos em Oro Waram (Wari’, Pakaa Nova, Txapakura) e as fronteiras fortes e fracas: ‘restrição’ e ‘permissão’ de regras fonológicas regulares*. Artigo apresentado no Congresso da ABRALIN, Natal (RN). (no prelo). 2013b.

APONTES, Selmo A. *Pronominalização em Oro Waram (Wari’, Pakaa Nova, Txapakura): formas condicionadas pelo TAM*. Artigo apresentado na Semana Internacional de Etnolinguística UNIR em Guajará-Mirim, RO. 2013c. (no prelo)

AVINO, Rita d’. *Introduzione a um corso di storia comparata delle lingue classiche*. Roma: Kappa, 1997.

BEKSTA, Pe. Casimiro. *Primeiras Letras para Oro Warí (Pakaa Nova – Oro Món, Oro Ramxién)*. Sagarana: Guajará-Mirim, 1982.

BJELJAC-BABIC. Seis mil lenguas, un patrimonio en peligro. *El correo de la Unesco*, abril 2000. Disponível em: <<http://unesdoc.unesco.org/images/0011/001194/119473s.pdf>>. Acesso em: 18 jul. 2011.

BLIKSTEIN, Izidoro. Indo-europeu, linguística e... racismo. *Revista USP*, n. 14, p. 104-110, jun.-ago. 1992. Disponível em: <<http://www.usp.br/revistausp/14/20-izidoro.pdf>>. Acesso em: 18 jul. 2011.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Instituto Brasileiro de Geografia e Estatística. *Censo Demográfico 2010*. Disponível em: <<http://www.ibge.gov.br/home/estatistica/populacao/censo2010>>. Acesso em: 10 jan. 2014.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Instituto Brasileiro de Geografia e Estatística. *Indígenas*. Disponível em: <<http://indigenas.ibge.gov.br>>. Acesso em: 10 jan. 2014.

BOLOGNINI, Carmen Zink; PAYER, Maria Onice. Línguas de imigrantes. *Ciência e cultura*, v. 57, n. 2, São Paulo, abril/junho 2005.

BONVINI, Emilio. “Classes d’accords” dans les langues négro-africaines. Un trait typologique du Niger-Congo. Exemples du kassim et du kimbundu. *Faits de langues*, v. 4, n. 8, p. 77-88, 1996.

- BONVINI, Emilio. Línguas africanas e português falado no Brasil. In: FIORIN, José Luiz; PETTER, Margarida (Orgs.). *África no Brasil*. São Paulo: Contexto, 2008. p. 15-62. N° **chamada (Fafich): 469.798 A258 2008**
- BRINTON, Daniel G. *The American Race: A linguistic classification and ethnographic description of the native tribes of north and South America*. Philadelphia: D. McKay, 1901
- BUCK, Carl D. *Dictionary of selected synonyms in the principal indo-european languages: a contribution of the history of ideas*. Chicago: University Press, 1949.
- BUNSE, Heinrich A. W. *Iniciação à filologia germânica*. Porto Alegre: Editora da UFRGS, 1983.
- CÂNDIDO, Gláucia V. *Descrição Morfossintática da língua Shanenáwa (Pano)*. 2003. Dissertação (Mestrado em Linguística). Campinas: UNICAMP, 2003. 264 f.
- CAGLIARI, Luiz Carlos. *Alfabetização & linguística*. São Paulo: Scipione, 2000. N° **chamada (Letras): 372.41 C131a 1996**
- CÂMARA Jr., Joaquim M. *Dicionário de filologia e gramática*. São Paulo/Rio de Janeiro/Fortaleza: Jozon, [s/d]. N° **chamada (Letras): 469.503 C172d 1973 R**
- CAMPANILE, Enrico. *La ricostruzione della cultura indoeuropea*. Pisa: Giardini, 1990.
- CARLIN, Eithne B. *A grammar of Trio*. Duisburger Arbeiten zur Sprach- und Kulturwissenschaft, v. 55. Frankfurt am Main: Peter Lang, 2004. (Europaischer Verlag der Wissenschaften).
- CASA NOVA, Vera. *Desertos*. Rio de Janeiro: 7Letras, 2004. N° **chamada (Letras): B869.141 C334c 2004**
- CASTRO, Yeda P. de. *Falares africanos na Bahia: um vocabulário afro-brasileiro*. Rio de Janeiro: Academia Brasileira de Letras/Topbooks, 2001. N° **chamada (Letras): 469.798 C355f 2001**
- CASTRO, Yeda P. de. *A língua mina-jeje no Brasil: um falar africano em Ouro Preto do século XVIII*. Belo Horizonte: Fundação João Pinheiro, 2002. N° **chamada (Letras): 417.2 C355l 2002**
- CAVALLI-SFORZA, Luigi. *Genes, povos e línguas*. São Paulo: Companhia das Letras, 2003. N° **chamada (Letras): 401 C377g.Pm 2003**
- CONSELHO DA EUROPA. European Charter for Regional or Minority Languages. Disponível em: <<http://conventions.coe.int/treaty/en/Treaties/html/148.htm>>. Acesso em: 4 dez. 2014.

CORBETT, Greville. *Gender*. Cambridge: CUP, 1991. N° **Chamada (Letras): 415 C789g 1991**

CROWTHER, Samuel. *Grammar of the Yoruba Language*. London: Seeleys, 1852.

CUNHA, Celso. *A questão da norma cultura brasileira*. Rio de Janeiro: Tempo Brasileiro, 1985. N° **chamada (Letras): 469.07 C972q**

CUNHA, Celso. *Gramática da língua portuguesa*. Rio de Janeiro: Fename, 1982. N° **Chamada (Letras): 469.5 C972g 1982**

CHATELAIN, Heli. *Gramática elementar do Kimbundu ou língua de Angola*. Geneva, [s/ed.], 1888-89. N° **de Chamada (Letras): 469.5 C972g 1982**

CHOMSKY, Noam. *Lectures on Government and Binding*. Dordrecht: Foris, 1981. N° **chamada (Letras): 415 C548l 1981**

D'ANS, André-Marcel *et al.* *Problemas de clasificación de Lenguas no-andinas en el su-este peruano*. Documento de trabalho n. 8. Centro de Investigación de Linguística Aplicada. Peru: Universidad Nacional de San Marcos, 1973.

D'ANS, André-Marcel. Reclassificación de las leguas pano y datos glotocronológicos para La etnohistoria de la Amazônia peruana. *Revista del Museo Nacional*, n. 39, p. 349-369, 1973.

DIAMONDS, Jared. *Armas, germes e aço*. Rio de Janeiro: Record, 2006. N° **chamada (Letras): 415 C548l 1981**

DIAS, Pedro. *Arte da lingua de Angola oeferecida a Virgem Senhora N. do Rosario, mãy, & Senhora dos mesmos Pretos, pelo P. Pedro Dias da Companhia de Jesu*. Lisboa: Officina de Miguel Deslandes, 1697.

DIRINGER, David. *A escrita*. Lisboa: Verbo, 1968. N° **chamada (Esc. Ciência da Inf.): 003 D598w.P 1968**

<http://dnghu.org/> Acesso em: 6 dez. 2014.

DURBIN, Marshall. A survey of the Carib language family. In: BASSO, Ellen. (Ed). *Carib-speaking indians: culture, society and language*. Tucson: University of Arizona Press, 1977. p. 23-38.

ECO, Umberto. *La ricerca della lingua perfetta nella cultura europea*. Roma-Bari: Laterza, 2002. N° **chamada (Letras): 410 E19r.Pa 2002**

ELIA, Sívio. *Preparação à linguística românica*. Rio de Janeiro: Ao Livro Técnico, 1979. N° **Chamada (Letras): 440 E42p 1979**

EVERETT, Daniel L.; KERN, Bárbara. *The Wari' language of western Brazil (Pacaas Novos)*. London: Routledge, [1997] 2006.

www.famsi.org/index.html Acesso em: 23 abr. 2010.

FERRAREZI JUNIOR, Celso. *A interface semântica/morfossintaxe*. 1998. Tese (Doutorado em Linguística). Guajará-Mirim: Campus da UNIR, 1998. 198 f.

FARACO, Carlos Alberto. *Linguística histórica: uma introdução ao estudo da história das línguas*. São Paulo: Parábola, 2005. N° **chamada (Letras): 808.8 F981 1991 v.78**

FERREIRA, Aurélio B. de H. *Novo dicionário da língua portuguesa*. Rio de Janeiro: Nova Fronteira, s/d. N° **chamada (Letras): 469.3 F383a 1999 R**

FRANÇA, Maria Cristina V. *Aspectos da fonologia lexical e pós-lexical da língua Oro Tawati' (Oro Win)*. 350 f. Tese (Doutorado em Linguística). Guajará-Mirim: Campus da UNIR, 2002.

FIORIN, José Luiz; PETTER, Margarida (Org.). *África no Brasil*. São Paulo: Contexto, 2008. N° **chamada (Fafich): 469.798 A258 2008**

FISCHER, Steven R. *A History of Writing*. Bury St Edmunds, Suffolk: St Edmundsbury Press, 2003.

GABBA, Emilio. Il latino come dialetto greco. In: *Roma arcaica: storia e storiografia*. Roma: Edizioni di Storia e Letteratura, 2000. p. 159-164. N° **chamada (Letras): 937 G112r 2000**

GAMKRELIDZE, Thomas V.; IVANOV, Vjačeslav V. Ivanov. *Indo-European and the Indo-Europeans*. Berlin/New York: Mouton de Gruyter, 1995. N° **chamada (Letras): 491 G747i.In 1995**

GIRARD, Victor. *Proto-Carib phonology*. 1971. 363 f. Tese (Doutorado em Linguística). University of California, Berkeley, 1971.

GOMES DE ARRUDA, (Dom) Roberto (s/d). *Dicionário Oro Waram-Português e Português-Oro Waram*. Guajará-Mirim: Diocese de Guajará-Mirim (Digitalizado).

GREENBERG, Joseph. *Language in the Americas*. Stanford: Stanford University Press, 1987.

GREENBERG, Joseph. *Studies in African Linguistic Classification*. Connecticut: Compass Publishing, 1955.

- GREENBERG, Joseph. *The Languages of Africa*. Bloomington: Indiana University, 1966.
- HALL, Katherine Lee. *The morphosyntax of discourse in De'kwana carib*. 1988. 428 f. Tese (Doutorado em Linguística). St. Louis: Washington University, 1988.
- HAWKINS, Robert. E. Waiwai. In: DERBYSHIRE, Desmond. C.; PULLUM, Geoffrey. K. (Eds.). *Handbook of Amazonian languages*. Berlin: Mouton de Gruyter, 1998. v. 4. p. 25-224.
- HAGÈGE, Claude. *Halte à la mort des langues*. Paris: Odile Jacob, 2000.
- HANGANU, Aurelia. Și iarăși despre... limba noastră. *Akademos*, Academia de Științe a Moldovei, n. 1, v. 12, p. 89-92, februarie 2009.
- HIGOUNET, Charles. *História concisa da escrita*. São Paulo: Parábola, 2003. N° *chamada (Letras): 411.09 H638e.Pm 2003*
- HIRT, Hermann. *Indogermanische Grammatik*. Heidelberg: Carl Winter's Universitätsbuchhandlung, 1927. N° *chamada (Letras): 491 I41 1969*
- HOLANDA, Sérgio Buarque de. *Raízes do Brasil*. Rio de Janeiro: José Olympio, 1936. N° *chamada (Letras): 981.01 H722r 2004*
- HORCADES, Carlos M. *A evolução da escrita*. Rio de Janeiro: Editora Senac Rio, 2004. N° *chamada (Fafich): 417.7 H811e 2004*
- HOUAISS, Antônio. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2001. N° *chamada (Letras): 469.3 H835d 2001 R*
- HULST, Harry van der. *The phonological structure of words: an introduction*. Cambridge: Cambridge University Press, 2001.
- INSTITUTO SOCIOAMBIENTAL. *Povos Indígenas no Brasil*. Disponível em: <<http://pib.socioambiental.org/pt/c/no-brasil-atual/linguas/introducao>>. Acesso em: 10 jan. 2014.
- JACKSON, Walter. S. A Wayana grammar. In: GRIMES, J. E. (Ed.). *Languages of the Guianas*. Norman: Summer Institute of Linguistics and University of Oklahoma Press, 1972. p. 47-77.
- JONES, William. The Third Anniversary Discourse. In *Works*, v. 1. [s.l.]: General Books LCC, 2009. p. 19-34.

- KATAMBA, Francis. Bantu Nominal Morphology. In: NURSE, Derek; PHILIPPSON, Gérard. *The Bantu Languages*. London/New York: Routledge, 2006. p. 103-120. N° *chamada (Letras): 496.39 B219 2003*
- KAUFMAN, Terrence K. The native languages of South America. In: MOSELEY, C.; ASHER, R. E. *Atlas of the World's Languages*. New York: Routledge, 1994. p. 46-76.
- KELLER, Rudi. *On language change: the invisible hand in language*. London/New York: Routledge, 1994. N° *chamada (Letras): 410 K29o 1994*
- KEY, Mary R. *Comparative Tacanan Phonology*. The Hague: Mouton, 1968.
- KOEHN, Edward; KOEHN, Sally. Apalai. In: DERBYSHIRE, Desmond. C.; PULLUM, Geoffrey. K. (Eds.). *Handbook of Amazonian languages*. Berlin: Mouton de Gruyter, 1986. v. 1. p. 33-127.
- LANES, Elder. J. *Aspectos da Mudança Linguística em um conjunto de línguas Amazônicas. As Línguas Pano. ?f. Tese (Doutorado em Linguística)*. UFRJ, 2005.
- LEITE, Serafim da Silva. *História da Companhia de Jesus no Brasil*. Lisboa: Portugália/Rio de Janeiro: Civilização Brasileira, 1938. N° *chamada (Fafich): 981.0311 L533h 1938-50*
- LEJEUNE, Michel; MARCH, Carlos A. R. *La posición del Latín en el Dominio Indoeuropeo*. Buenos Aires: Coni, 1949.
- LEWIS, M. P. (Ed.). *Ethnologue: Languages of the World*. Dallas: SIL, 2009. Disponível em: <<http://www.ethnologue.com/>> Acesso em: 18 jul. 2011.
- LOOS, Eugene. Pano. In: DIXON, Robert M. W.; AIKHENVALD, Alexandra Y. (Eds.). *The Amazonian Languages*. Cambridge: CUP, 1999, p. 226-250.
- LOOS, Eugene. Rasgos sintáctico-fonémicos en la historia lingüística de los idiomas de la familia pano. *Linguística e Indigenismo moderno de América*, v. 5, p. 181-184, 1975.
- MACDONELL, Ronald B. 1994. 594 f. *La phonologie du Makuxi, langue caribe: une analyse fonctionnelle*. Dissertação (Mestrado em Linguística). Québec: Université Laval, 1994.
- MACHADO FILHO, Aires da Mata. *O negro e o garimpo em Minas Gerais*. Rio de Janeiro: José Olympio, 1943. N° *chamada (Fafich): 981.00821 D637 v.42*
- MAEDA, Cláudia T. *Descrição Preliminar do Oro Eo: um caso de sílaba embutida no onset*. 2000. 129 f. Dissertação (Mestrado em Linguística). Guajará-Mirim: Campus da UNIR, 2000.
- MARTINS, Wilson. *A palavra escrita*. São Paulo: Ática, 2002. N° *Chamada (Letras): 411.09 M386p 2001*

MASON, John A. The languages of South American Indians. In: STEWARD, Julian. H. (Ed.) *Handbook of South American Indians*, v. 6. Washington, DC: Smithsonian Institution. p. 157-317, 1950.

MATTEI-MULLER, Marie-Claude. *Diccionario ilustrado panare-español, español-panare: un aporte al estudio de los Panares-E'ñepa*. Caracas: Comisión Quinto Centenario, Gráficas Armitano, 1994.

MEILLET, Antoine. *Les langues dans l'Europe nouvelle*. Paris: Payot, 1928.

MEILLET, Auguste. *Linguistique historique et linguistique générale*. Leiden: E. J. Brill, 1948. N° *chamada (Letras): 410 M513l 1982*

MEIRA, Sérgio. *A grammar of Tiriyo*. 708 f. 1999. Tese (Doutorado em Linguística). Houston: Rice University, 1999.

MEIRA, Sérgio; FRANCHETTO, Bruna. The southern Cariban languages and the Cariban family. *International Journal of American Linguistics*, v. 71, p. 127-192, 2005.

MEIRELES, Denise M. *Os Pakaas-Novos*. 525 f. 1986. Dissertação (Mestrado em Antropologia). Brasília: UNB, 1986.

MOORE, Denny. Endangered Languages of Lowland Tropical South America. In: BENZINGER, Matthias (Org.). *Linguistic Diversity Endangered*. Berlin: De Gruyter Incorporated, 2007. p. 29-58.

MOORE, Denny. Endangered languages of Lowland Tropical South America. In: BREZINGER, Matthias. *Language Diversity Endangered*. Berlin: Mouton de Gruyter, 2007a. p. 29-58.

MOORE, Denny. Línguas indígenas: situação atual, levantamento e registro. *Patrimônio - Revista Eletrônica do IPHA*, n. 6, jan./fev. 2007b. páginas? Disponível em: <http://www.institutobuzios.org.br/documentos/Denny%20Moore_LINGUAS%20IND%CDG_ENAS%20SITUA%C7%C3O%20E%20REGISTRO.pdf>. Acesso em: 29 dez. 2014.

MOORE, Denny; GABAS, Nilson. O Futuro das Línguas Indígenas Brasileiras. In: FORLINE, Louis; VIEIRA, Ima; MURRIETA, Rui (Orgs.). *Amazônia além dos 500 Anos*. Belém: Museu Paraense Emílio Goeldi, 2006. p. 433-454.

MOORE, Denny; STORTO, Luciana. *Linguística Indígena no Brasil*. 1992. Ms.

MOSCATI, Sabatino *et al.* (Coord.). *I celti*. Milano: Bompiani, 1991.

MÜLLER, Diocelma Maria. *Aspectos da Fonética, da Fonêmica e do Léxico da Língua Moré (Família Txapakura)*. 156 f. 1995. Dissertação (Mestrado em Linguística). Florianópolis: UFSC, 1995.

MUSEU PARAENSE EMÍLIO GOELDI. *Línguas Indígenas da Amazônia*. Disponível em: <http://saturno.museu-goeldi.br/lingmpeg/portal/?page_id=259>. Acesso em: 10 jan. 2014.

NIKOLSKI, W.; JAKOWLEW, N. *Warum die Menschen in verschiedenen Sprachen sprechen*. Berlin: Verlag der Sowjetischen Militärverwaltung, 1947.

NIMUENDAJU, Curt; BENTES, E. H. do Valle. Document sur quelques langues peu connues de l'Amazonie. *Journal de la Société des Américanistes*, Paris, t. 14-15, 1922, p. 215-222.

NOLL, Volker; DIETRICH, Wolf. *O português e o tupi no Brasil*. São Paulo: Contexto, 2010.

NURSE, Derek; PHILIPPSON, Gérard. *The Bantu Languages*. London/New York: Routledge, 2006.

www.omniglot.com/writing/definition.htm Acesso em: 23 abr. 2010.

ONG, Walter. *Oralidade e cultura escrita*. Campinas: Papyrus, 1998. N° **chamada (Letras): 302.2 O58o.Pd 1998**

OSTLER, Rosemarie. Searching for the first words. *Verbatim: The Language Quaterly*, v. 28, n. 4, p. 1-4, Winter 2003. Disponível em: <www.verbatimmag.com/28_4.pdf>. Acesso em: 18 jul. 2011.

www.proel.org Acesso em: 29 maio 2014.

PACHÊCO, Frantomé B. *Morfossintaxe do verbo Ikpeng (Karib)*. 2001. 303 f. Tese (Doutorado em Linguística). Campinas: UNICAMP, 2001.

PAULA, Aldir. S. de. *A língua dos índios Yawanawá do Acre*. 2004. f. Tese (Doutorado em ?). Campinas: UNICAMP, 2004.

PAYNE, Doris L. Noun Classification in the Western Amazon. In: *Languages Sciences*, v. 9, n. 1, 1987. p. 21-44.

QUEIROZ, Sônia. *Pé preto no barro branco: a língua dos negros da Tabatinga*. Belo Horizonte: Editora UFMG, 1998. N° **chamada (Letras): 417.2 Q3p 1998**

QUINT, Nicolas. A realização do sujeito em português do Brasil: deriva *versus* criouliização. In: FIORIN, José Luiz; PETTER, Margarida (Org.). *África no Brasil*. São Paulo: Contexto, 2008. p. 75-88. N° **de Chamada (Fafich): 469.798 A258 2008**

RAMAT, Anna Giacalone; RAMAT, Paolo (Cur.). *Le lingue indoeuropee*. Bologna: Il Mulino, 1993.

RAMIREZ, Henri. Etnônimos e topônimos no Madeira (séculos XVI-XX): um sem-número de equívocos. *Revista Brasileira de Linguística Antropológica*. v. 2, n. 2. dez. de 2010. p. 179-223.

RIBEIRO, Darci. Culturas e Línguas Indígenas do Brasil. *Educação e Ciências Sociais*. Rio de Janeiro: v. 2, n. 6, nov. de 1957. p. 4-102.

RIBEIRO, Eduardo R. Macro-Jê. In: BROWN, Keith. *Encyclopedia of Language and Linguistics*. v. 7, Oxford: Elsevier, 2006. p. 422-426.

RODRÍGUEZ ADRADOS, Francisco. *Linguística indoeuropea*. Madrid: Gredos, 1975. Nº **chamada (Letras): 491 A2421 1975**

RODRIGUES DURAN, Ires. *Descrição fonológica e lexical do dialeto Kaw Tayo da Língua Moré*. 181f. 2000. Dissertação (Mestrado em Linguística) Guajará-Mirim: Campus da UNIR, 2000.

RODRIGUES, Aryon D. Classification of Tupí-Guaraní. *International Journal of American Linguistics*, v. 24, p. 231-234, 1958.

RODRIGUES, Aryon D. Tarefas da Linguística no Brasil. *Revista Brasileira de Linguística Teórica e Aplicada*, v. 1, n. 1, p. 4-15, 1966.

RODRIGUES, Aryon D. *Línguas Brasileiras: para conhecimento das línguas indígenas*. São Paulo: Loyola, 2002.

RODRIGUES, Aryon D. *Línguas Brasileiras: para conhecimento das línguas indígenas*. São Paulo: Loyola, 1986.

RODRIGUES, Aryon D. Macro-Jê. In: DIXON, Robert M. W.; AIKHENVALD, Alexandra Y. (Orgs.). *The amazonian languages*. Cambridge: Cambridge University Press, 1999. p. 164-206.

RODRIGUES, Aryon D.; CABRAL, Ana Suelly A. Revendo a classificação interna da família tupi-guarani. In: ____ (Orgs.). *Línguas Indígenas Brasileiras. Fonologia, gramática e história. Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da Anpoll*. Belém: UFPA, 2002. p. 327-337.

RODRIGUES, Aryon D. Sobre as línguas indígenas e sua pesquisa no Brasil. *Ciência e Cultura*, Jun. 2005, v. 57, n. 2, p. 35-38.

RONDON, Gen. Cândido M. S.; FARIA, Dr. João B. de. *Glossário Geral das tribos silvícolas de Mato-Grosso e outras da Amazônia e do Norte do Brasil*. Tomo I. Rio de Janeiro: Imprensa Nacional, 1948.

RYDEN, Stig. *Los índios Moré: notas etnográficas*. Bolivia, La Paz: Ministerio de Educación - Departamento de Arqueologia, Etnografía y Folklore, 1958.

RUGY, François de. À l'Assemblée, il est interdit de parler breton... Disponível em: <<http://auxmarchesdupalaisbourbon.blogs.liberation.fr/2008/2008/05/a-lassemble-il.html>>.

Acesso em: 18 jul. 2011.

SAUSSURE, Ferdinand de. *Curso de linguística geral*. São Paulo: Cultrix, s/d. N° *chamada (Letras): 410 S259c.Pc 1989*

SAUSSURE, Ferdinand de. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Leipsick: B. G. Teubner, 1879. Disponível em: <http://www.revue-texto.net/Saussure/De_Saussure/Memoire/N0072920_PDF_1_306.pdf>. Acesso em: 18 jul. 2011.

SAEZ, Oscar Calávia. *O nome e o tempo em Yaminawa: etnologia e história dos Yaminawa do rio Acre*. São Paulo: Editora UNESP: ISA; Rio de Janeiro: NUTI, 2006.

SHEL. *Las lenguas Pano y sy reconstrucción*. Estudios Panos III, Série Linguística Peruana, 10, Peru. ILV, 1975.

SILVA, Oziel Marques da. *A arquitetura das representações fonológicas dos empréstimos portugueses em Oro Mon (Wari' Norte, Rondônia)*. 2003. 89 f. Dissertação (Mestrado em Linguística). Guajará-Mirim-RO: Campus da UNIR, 2003.

SOUSA, Maria de Fátima Lima. *Dicionário da Língua Wari': Oro Mon – Português*. 2009. 140 f. Dissertação (Mestrado em Linguística). Guajará-Mirim: Campus da UNIR, 2009.

SOUZA, Shirley D. *Alguns aspectos morfológicos da língua arara (Karíb)*. 1993. 76 f. Dissertação (Mestrado em Linguística). Brasília: Universidade de Brasília, 1993.

SOUZA, Tânia C. *Discurso e oralidade: um estudo em língua indígena*. 1994. 398 f. Tese (Doutorado em Linguística). Campinas: UNICAMP, 1994.

SPANGUERO FERREIRA, R. *Estudo lexical da língua Matís: subsídios para um dicionário bilíngue*. 2003. 211 f. Tese (Doutorado em Linguística). Campinas: UNICAMP, 2003.

STEINEN, Karl von den. *Die Bakairi-Sprache: Wörterverzeichnis, Sätze, Sagen, Grammatik. Mit Beiträgen zu einer Lautlehre der karäibischen Grundsprache*. Leipzig: Koehler, 1892.

STEINEN, Karl von den. *Durch Central-Brasilien: Expedition zur Erforschung des Schingú im Jahre 1884*. Leipzig: F. A. Brockhaus, 1886.

TAVARES, P. S. *A grammar of Wayâna*. 2005. 464 f. Tese (Doutorado em Linguística). Houston: Rice University, 2005.

<http://tipografos.net>. Acesso em: 23 abr. 2010.

VALENZUELA, Pilar M. *Transitivity in Shipibo-Konibo Grammar*. 1129 f. 2003. Tese (Doutorado em Linguística). Oregon: USA, 2003.

VIDOS, B. E. *Manual de lingüística románica*. Traducción de Francisco B. de Moll. Madrid: Aguilar, 1973. N° *chamada (Letras): 440 V657h.Em 1968*

VILLAR LIÉBANA, Francisco. *Lenguas y pueblos indoeuropeos*. Madrid: Istmo, 1971. N° *chamada (Letras): 491 V719e*

VILLAR LIÉBANA, Francisco. *Origen de la flexión nominal indoeuropea*. Madrid: Instituto Antonio de Nebrija, 1974. N° *chamada (Letras): 491 V719o 1974*

VILLAR LIÉBANA, Francisco. *Los indoeuropeos y los Orígenes de Europa, Lenguage e historia*. Madrid: Gredos, 1991.

VILAÇA, Aparecida. *Comendo como gente: formas do canibalismo Wari' (Pakaa Nova)*. Rio de Janeiro: Editora UFRJ, 1992.

VOGT, Carlos; FRY, Peter. *A África no Brasil*. Campinas: Unicamp, 1996.

WATKINS, Calvert. *The American Heritage Dictionary of Indo-European Roots*. Boston: Houghton Mifflin, 2000.

YITZJAKI, Shlomó. *El Pentateuco con el comentario de Rabí Shlomó Yitzjaki (Rashi)*. Traducción directa al castellano del original hebreo por Enrique Jaime Zadoff (El Pentateuco) y Jaime Barilko (Rashi). Buenos Aires: Editorial Yehuda, [s/d].