

## A RECONSTRUÇÃO COMPARATIVA

Texto traduzido e adaptado de  
“Chapter 2. Comparative Reconstruction” (pp. 17-36) em  
Robert J JEFFERS e Ilse LEHISTE, *Principles and Methods for Historical Linguistics*.  
Cambridge, MASS e Londres, Inglaterra: The MIT Press, 1979 (5ª impressão de 1992).

O método comparativo é baseado em dois pressupostos fundamentais. O primeiro argumento, poderíamos denominar a *hipótese do parentesco* e o segundo a *hipótese da regularidade*. A hipótese do parentesco procura explicar semelhanças evidentes entre vocábulos que pertencem a línguas ou dialetos diferentes por meio da suposição de que esses dialetos e línguas são parentes. A teoria pressupõe que as línguas e os dialetos em consideração descendem de um antecessor comum ou *protolíngua*. A hipótese da regularidade possibilita que essa protolíngua seja reconstruída ao supor que a mudança fonológica é um fenômeno regular. Essa teoria conjectura que, caso mudar, cada som num dado dialeto será modificado da mesma maneira em cada instância nas mesmas circunstâncias. O método comparativo consiste em examinar palavras com significados parecidos em línguas suspeitadas de descender de uma protolíngua comum, na esperança de descobrir *correspondências fonológicas* e de reconstruir a protolíngua.

### Como estabelecer as correspondências fonológicas

O procedimento envolve o exame dos sons numa determinada posição dentro de certo morfema. Por exemplo, comparam-se as consoantes iniciais num conjunto de vocábulos para os quais existe a suspeita de as palavras serem cognatas. Os *cognatos* são palavras que descendem do mesmo vocábulo na protolíngua. É comum que itens cognatos exibam semelhanças tanto na sua forma quanto no seu significado. Após examinar todos os sons situados em posições análogas e quando todas as correspondências fonológicas estiverem identificadas, é possível progredir à reconstrução da forma da palavra na protolíngua.

De modo a exemplificar, consideremos uma eventual reconstrução da palavra indo-europeia que significa “nuvem” baseada na comparação dos vocábulos glosados “nuvem” em três línguas que suporemos ser aparentadas: sânscrito (sâns.) नभस (*nábas* /'na.b<sup>h</sup>as/), grego antigo (gr.) νέφος (*néphos* /'ne.p<sup>h</sup>os/) e o eslavônico eclesiástico antigo (EEA) нѣбо (*nebo* /'ne.bo/). “Indo-europeu” (IE) é o termo que utilizamos para caracterizar uma grande família de línguas faladas em tempos antigos e modernos na Europa e no oeste da Ásia. Trata-se da família linguística a qual pertence o português. A protolíngua de que as línguas IE derivam é denominada tipicamente o *protoindo-europeu* (PIE). Observamos semelhanças na forma e identidade do significado nesse conjunto de palavras, e conjecturamos que as mesmas constituem um conjunto de cognatos – um conjunto de morfemas ou palavras em línguas atualmente diferentes derivadas de um único “morfema pai” ou “palavra mãe”. Os *conjuntos de correspondência* para essas três palavras são exibidos abaixo em (2.1). Os conjuntos correspondentes identificados para a palavra “nuvem” nessas três línguas se repetem em muitos outros morfemas dessas línguas.

(2.1)

sânscrito	grego clássico	eslavônico eclesiástico antigo
n	n	n
a	e	e
bh	ph	b
a	o	o
s	s	...

Na base de correspondências como essas, procura-se estabelecer uma forma reconstruída da qual todos os sons nas línguas “filha” poderiam ser derivados por meio de mudanças fonológicas plausíveis. O termo “*língua filha*” é utilizado para designar cada uma de um número qualquer de línguas aparentadas como descendente histórica ou continuação de alguma língua anterior. A relação cognática que existe entre um grupo de “línguas filhas” é caracterizada pelo termo “*língua irmã*”. A reconstrução de *n* como a consoante inicial é simples no presente caso, já que todas as três línguas exibem o mesmo reflexo. O termo *reflexo* é utilizado nesse contexto para se referir a um som que ocupa determinada posição em certo morfema, o qual parece ser a continuação de um som anterior que ocupava a mesma posição no mesmo morfema.

A reconstrução da vogal na primeira sílaba é menos direta. O /a/ do sânscrito corresponde a /e/ tanto no grego antigo como no eslavônico eclesiástico antigo. Já que duas das línguas filhas se opõem contra uma, parece mais provável que uma das línguas tenha inovado e duas tenham mantido o estado original. Além disso, embora não precisemos entrar em muito detalhe neste ponto, um desenvolvimento de /e/ para /a/ neste contexto é um tanto mais plausível por motivos fonéticos.

Na reconstrução da consoante medial, notamos que cada uma das três línguas profere um reflexo diferente. No entanto, cada reflexo contém uma oclusiva bilabial. O sânscrito e o eslavônico eclesiástico antigo compartilham o traço de vozeamento, enquanto o sânscrito e o grego antigo compartilham o traço da aspiração. Em virtude do segmento que aparece no sânscrito exibir um traço em comum com cada uma das suas línguas irmã, *bh* é geralmente reconstruído para esse conjunto correspondente. A escolha do símbolo *bh* é de um interesse especial quanto à natureza das formas reconstruídas. Embora esse símbolo seja utilizado extensivamente, a forma fonética verdadeira do segmento reconstruído é um tema de controvérsia entre os especialistas no indo-europeu. O símbolo é meramente uma fórmula que serve para representar a correspondência observada. A realidade fonética atrás das formas reconstruídas é muitas vezes difícil de estabelecer. O que deve ser enfatizado é que as formas reconstruídas simbolizam correspondências e, como tal, não representam qualquer tentativa definitiva de estabelecer as formas fonéticas das palavras reconstruídas. Um segmento reconstruído é uma unidade estrutural do sistema fonológico ancestral.

Como foi o caso da primeira vogal nessa palavra, no que diz respeito à vogal da segunda sílaba, o grego antigo e o eslavônico eclesiástico antigo concordam nos seus reflexos, ao contrário do sânscrito. Por motivos parecidos aos apresentados acima, reconstruiríamos *o* para a língua ancestral. Vale ressaltar de novo, porém, que a qualidade fonética exata das vogais reconstruídas é de menor relevância do que o fato de o sânscrito ter reestruturado seu sistema fonológico. Onde a língua mãe apresentava duas vogais distintas, o sânscrito possui apenas uma vogal. A alternativa seria de supormos que o sânscrito reflete a situação arcaica e que o grego antigo e o eslavo eclesiástico antigo inovaram. Entretanto, se um segmento da língua ancestral exibir mais do que um reflexo numa dada língua filha sua e, se a hipótese da regularidade for correta, esperaríamos poder estabelecer o contexto fonético em que cada reflexo se desenvolveu. Nenhum contexto plausível se apresenta nos dados, neste caso.

O sânscrito e o grego antigo de correspondem ao exibir uma consoante final, *s*. Não há tal consoante final na palavra EEA. Em casos como esse, é melhor reconstruir a palavra com uma

consoante final e supor que esse *s* foi eliminado no eslavo eclesiástico antigo, pelo menos em posição final das palavras. A forma reconstruída da palavra PIE para “nuvem” se apresentaria como *\*nebhos*. O asterisco é usado para indicar o fato de estarmos a tratar de uma palavra não atestada, reconstruída antes de ser uma palavra que ocorra verdadeiramente numa língua que conheçamos por meio de registros escritos, ou numa língua que seja usada atualmente por alguma comunidade linguística.

Podemos ver, portanto, que, para cada correspondência que o linguista consiga estabelecer, um segmento pode ser postulado (pelo menos de modo temporário) para a língua ancestral. Se for possível mostrar, entretanto, que dois ou mais conjuntos correspondentes ocorrem em ambientes contrastantes, estamos justificados ao reconstruir um único segmento na língua mãe para conjuntos contrastantes. O exemplo abaixo, baseado em Hoenigswald (1950), mostra o procedimento seguido pelo linguista ao reconstruir os segmentos do sistema fonológico ancestral.

(2.2)

	<b>sânscrito</b>	<b>gótico</b>	<b>glossa portuguesa</b>
1.	/t/ असति ( <i>ásti</i> )	/t/ IΣT ( <i>ist</i> )	“é”
2.	/t/ पितर ( <i>pitár</i> )	/d/ FAΔAR ( <i>fadar</i> )	“pai”
3.	/t/ भ्रातर ( <i>bhrátar</i> )	/θ/ BPOΨAR ( <i>broþar</i> )	“irmão”
4.	/d/ देही ( <i>déhī</i> )	/d/ ΔΙΓΑΗ ( <i>digan</i> )	“parede”, “sovar”
5.	/d/ वेद ( <i>véda</i> )	/t/ YAIT ( <i>wait</i> )	“saber”
6.	/dh/ मध्यस ( <i>mádhyas</i> )	/d/ MIΔGIΣ ( <i>midjis</i> )	“meio”

Determinados detalhes dos dados pertinentes do sânscrito e do gótico foram desconsiderados aqui para facilitar nossa exposição do método. Não obstante, o exemplo (2.2) dá, de fato, um quadro razoavelmente certo da reconstrução das consoantes dentais PIE na base dos reflexos índicos e germânicos. Embora só um par de cognatos tenha sido apresentado para cada correspondência, é implícito que esses itens cognatos são representativos das correspondências que ocorrem de forma regular pelo léxico das duas línguas. O investigador pode estabelecer provisoriamente seis segmentos na protolíngua, cada um para representar cada conjunto correspondente. A seguir, o pesquisador precisa considerar o ambiente fonético em que essas correspondências ocorrem para verificar a presença de contraste.

A correspondência 2 (sânsc., /t/ → gót., /d/) ocorre apenas em ambientes vozeados e não é precedida pelo acento nas formas do sânscrito. Em contraste, a correspondência número 3 (sânsc., /t/ → gót., /θ/) ocorre em todas as posições das quais a correspondência 2 está excluída. Consequentemente, um único segmento *\*/t/* pode ser estabelecido para as correspondências 2 e 3. Por ocorrer em quase todos os ambientes, a correspondência 5 (sânsc., /d/ → gót., /t/) é bastante provável de refletir um segmento na língua mãe, digamos *\*/d/*. Algumas poucas exceções transparentes existem, como, por exemplo, as situações em que o /d/ sânscrito sofre desvozeamento devido a uma regra de assimilação de vozeamento. A correspondência 4 (sânsc., /d/ → gót., /d/) ocorre apenas quando uma consoante aspirada se manifesta na sílaba seguinte das formas sânscritas, enquanto pode-se afirmar que o número 6 (sânsc., /dh/ → gót., /d/) ocorre onde quer que o número 4 não se aplique. Deste modo, *\*/dh/* pode ser estabelecido como o protótipo para as correspondências 4 e 6. A correspondência 1 é mais difícil de se resolver, mas, para as nossas finalidades atuais, basta dizermos (outros dados relevantes existem) que

essa correspondência ocorre apenas depois de *s*, enquanto as correspondências 2 e 3 não aparecem nesse contexto. Portanto, já que o número 1 está numa distribuição complementar com os números 2 e 3, podemos reconstruir aquela correspondência como mais um dos reflexos de *\*t/*.

Este breve exercício em comparação histórica revela que, embora as consoantes dentais do sânscrito e do gótico exibam seis conjuntos correspondentes diferentes, o linguista precisa reconstruir somente três segmentos para o sistema fonológico ancestral. Além disso, o fato de que um único segmento possa sofrer mudanças diferentes em ambientes fonéticos diferentes se demonstra claramente pelo desenvolvimento de *\*t* PIE no gótico (onde observamos três reflexos) e da evolução de *\*dh* do sânscrito (que exibe dois reflexos).

De modo a treinar mais a reconstrução, considere os conjuntos de cognatos em (2.3) abaixo, que se baseiam em dados parciais selecionados tirados de três línguas balto-fineses da família fino-ugriana. Nesta reconstrução do protobalto-finês, uma consoante indica duração (ou seja, geminação) e a duração vocálica presente no livônio é expressa pelo uso do símbolo < : >; <ä> é uma vogal baixa anterior não arredondada, e <ü> é uma vogal alta anterior arredondada; <j> representa o /i/ assilábico e, no livônio, <l'> e <r'> indicam [ɫ, lʲ] e [rʲ], respectivamente; no estoniano, <g> e <d> são consoantes brandas *surdas*; <g> e <d> no livônio são sonoros.

(2.3)

	livônio	finlandês	estoniano	glossa
1.	<i>säv</i>	<i>savi</i>	<i>savi</i>	“barro”, “argila”
2.	<i>tämm</i>	<i>tammi</i>	<i>tamm</i>	“carvalho”
3.	<i>säpp</i>	<i>sappi</i>	<i>sapp</i>	“bólis”
4.	<i>lüm</i>	<i>lumi</i>	<i>lumi</i>	“neve”
5.	<i>o:da</i>	<i>hauta</i>	<i>haud</i>	“sepultura”
6.	<i>umal</i>	<i>humala</i>	<i>humal</i>	“lúpulo”
7.	<i>ja:lga</i>	<i>jalka</i>	<i>jalg</i>	“pé”
8.	<i>ne:l'a</i>	<i>neljä</i>	<i>neli</i>	“quatro”
9.	<i>ä:rga</i>	<i>härkä</i>	<i>härg</i>	“boi”
10.	<i>o:r'a</i>	<i>harja</i>	<i>hari</i>	“escova”, “pincel”

Para o conjunto de cognatos 1, precisamos identificar a qualidade da primeira vogal, e se o *i* final é original. Todos os demais traços dos cognatos são correspondentes. A vogal anterior da palavra *säv* no livoniano pode ser explicada como o resultado de um processo de assimilação à distância (metafonia/*umlaut*), pelo qual uma vogal não anterior (ou seja, central ou posterior) avança quando é seguida por uma vogal anterior na sílaba seguinte. Obviamente, a mudança fonológica sugerida supõe que o *i* final é original e que a palavra reconstruída é *\*savi*. O surgimento espontâneo de vogais finais é pouco frequente. Por motivos independentes, anteciparíamos a perda do *i* final no livônio para explicar o conjunto correspondente *0 : i : i / \_\_#*. Observe também que as formas 2, 3 e 4 exibem o mesmo conjunto correspondente. Além disso, as formas 2, 3 e 4 também apoiam nossa hipótese sobre o fenômeno da metafonia livoniana.

Se reconstruirmos *\*tammi* para 2, a forma livônio se explica pelas mudanças fonológicas já estabelecidas. A perda do *\*i* final no estoniano precisa de explicação, no entanto. A várias formas estonianas falta-lhes uma vogal nas posições onde tal vogal ocorre no finlandês e, às vezes, no livônio. Em todas as formas, com a exceção de 1 e 4, o estoniano não exibe nenhuma vogal em final de palavra, onde uma vogal ocorre no finlandês. As formas 1 e 4 diferem de todas as outras formas exibidas no

exemplo (2.3) na medida em que aquelas são as únicas formas em que a vogal final é precedida pela sequência CVC\_\_\_. Em todas as demais formas, a vogal final é antecedida ou por uma consoante longa (2, 3), por um grupo de consoantes (7, 8, 9, 10), por duas sílabas (6), ou por uma vogal na sílaba seguinte que não é breve (o ditongo em 5). Assim, poderíamos propor a seguinte regra como uma tentativa de explicar o comportamento do estoniano:

(2.4)

$$V > 0 / \left\{ \begin{array}{c} CC \\ V:C \\ CVCV \end{array} \right\} \_ \#.$$

Se reconstruirmos *\*sappi* para 3 e *\*lumi* para 4, nenhuma mudança fonológica nova precisa ser proposta.

Na quinta forma, o finlandês e o estoniano concordam em exibir uma consoante desvozeada branda média (apesar da <d> da ortografia estoniana), enquanto a forma livoniana apresenta uma consoante vozeada. Uma correspondência parecida se nota no caso da oclusiva velar em 7, o que sugere uma mudança no livônio por meio da qual as oclusivas surdas (ou, na base desses dados, pelo menos *t* e *k*) se sonorizam quando forem circundadas por segmentos vozeados. Os dados só fornecem evidência desse vozeamento médio no ambiente dos soantes (S) (vogais, líquidas, glides), pois, poderíamos caracterizar a mudança fonológica livoniana da seguinte maneira: *t, k > d, g / S\_\_S*. A correspondência *0 : h : h / #\_\_* em 5, 6 e 10 indica a perda de /h/ inicial no livônio e é provável que a presença de /o/ em *o:da* na mesma língua resulte de um processo de monotongação de *au*, que fica inalterado no finlandês e no estoniano. Além disso, o segmento *\*a* do protobalto-finês deve desaparecer em posição final no livônio nas palavras com mais de duas sílabas, de modo a explicar o número 6 livoniano, *umal* “lúpulo”.

Em 7, o livônio apresenta uma vogal longa onde os demais idiomas não exibem tal forma. O mesmo fenômeno ocorre em 8, 9 e 10, de modo que formulamos a correspondência geral  $\underline{V} : V : V$ . Em cada caso, a vogal em questão é seguida por um líquido (*l* ou *r*). Consequentemente, uma mudança no livoniano do tipo  $V > \underline{V} / \_ L$ , parece razoável. O número 7 seria reconstruído como *\*jalka*.

Em 8 e 9, percebemos o conjunto correspondente  $a : \ddot{a} : 0 / \_ \#$ . Já tivemos ocasião de explicar a perda das vogais finais no estoniano, mas a diferença na qualidade vocálica exibida pelo finlandês e o livoniano pede uma explicação. Observe que todas as línguas concordam em apresentar vogais anteriores na primeira sílaba (*e* e *ä*). Um processo de assimilação *progressiva* da qualidade vocálica se propõe para o finlandês, ou seja,  $a > \ddot{a} / V^{ANT} \_$ . Compare o processo parecido de assimilação *regressiva* (metafonia/*umlaut*) que propomos acima para o livônio.

Em *neli* do estoniano, deparamos com uma vocalização inesperada de *j* depois da consoante subsequente à perda da vogal final, ou seja,  $j > i / C \_ \#$ . *Hari* em estoniano (10) exibe a mesma evolução. Os números 8 e 10 em livônio se distinguem dos demais pela presença de líquidas palatalizadas. O glide assilábico *j* em finlandês, que supusemos ter passado a ser silábico no estoniano, é certamente a forma a ser reconstruída no protobalto-finês e é eliminado aparentemente no livônio após ter condicionado uma palatalização das líquidas, em que  $l, r > l', r' / \_ j$  e  $j > 0 / L' \_$ . Por último, a vogal da primeira sílaba no exemplo 10 parece ser *a*. Sugerimos já uma mudança fonológica no livônio em que  $V > \underline{V} / \_ L$ , o que explica a duração em *o:r'a*, e, embora a explicação histórica

verdadeira seja muito mais complexa, com base nos dados apresentados aqui, podemos propor, de modo experimental, um desenvolvimento pelo qual *o*: < *a* / \_\_ \$ (ou seja, numa sílaba *aberta*).

As formas reconstruídas do protobalto-finês são as seguintes: \**savi*, \**tammi*, \**sappi*, \**lumi*, \**hauta*, \**humala*, \**jalka*, \**nelja*, \**härka*, \**harja*. As mudanças fonológicas que afetaram cada língua são as seguintes:

(2.5)

**livônio**

1. V [-ant.] > V [+ant.] / \_\_ Ci
2. i > 0 / \_\_ #
3. C [-voz.] > C [+voz.] / soante \_\_ soante
4. h > 0
5. au > o
6. a > 0 / CVCVC \_\_
7. r, l > r', l' / \_\_ j
8. j > 0 / r', l' \_\_ V
9. V > V > \_\_ r, l
10. a: > o: / \_\_ CV

**finlandês**

- a > ä / V [+ant.] C \_\_

**estoniano**

1. V > 0 / CC, CVCVC, VC \_\_ #
2. j > i / C \_\_ #

Observe que, no livônio, a regra 1 deve preceder a regra 2, tal como a regra 7 precisa anteceder a regra 8, respectivamente, quanto à sua cronologia relativa. No estoniano, a mudança 1 ocorre necessariamente antes da mudança 2.

Até este ponto, temos prestado pouca atenção à questão do caráter fonético dos segmentos que temos reconstruído nos sistemas fonológicos das protolínguas. Entretanto, a precisão fonética está ficando mais viável graças a avanços na teoria fonológica. De modo a demonstrarmos algumas das possibilidades de maior exatidão no detalhamento fonético das reconstruções, voltemos à questão da reconstrução do sistema de consoantes PIE. No exemplo (2.2), uma reconstrução parcial das consoantes oclusivas foi realizada. Chegamos a propor uma série de consoantes dentais que incluía contrastes entre um segmento surdo (*t*), outro sonoro (*d*) e um terceiro segmento sonoro aspirado (*dh*). Essa série corresponde àquela que era reconstruída tradicionalmente para o PIE. Séries semelhantes de oclusivas labiais, palatovelares e labiovelares eram reconstruídas, produzindo o seguinte subsistema de oclusivas na protolíngua:

(2.6)

<b>Surdo</b>	<b>Sonoro</b>	<b>Aspirado</b>
<i>p</i>	? <i>b</i>	<i>b<sup>h</sup></i>
<i>t</i>	<i>d</i>	<i>d<sup>h</sup></i>
<i>k</i>	<i>g</i>	<i>g<sup>h</sup></i>
<i>k<sup>w</sup></i>	<i>g<sup>w</sup></i>	<i>g<sup>wh</sup></i>

O sistema parecido ao exibido acima em (2.6) foi tomado como certo pela maioria dos estudiosos do indo-europeu desde o século XIX. Na década de 1970, porém, a natureza fonética do sistema tradicional foi posta seriamente em dúvida como consequência de refinamentos numa teoria em vias de desenvolvimento acerca dos universais fonológicos. As maiores dificuldades surgem dos seguintes

fatos: (1) nenhuma língua em que exista um contraste de vozeamento, exhibe uma série de consoantes aspiradas sonoras na ausência de uma série de consoantes aspiradas surdas. Se apenas uma série de consoantes aspiradas estiver presente, essa série é universalmente a surda. (2) Existe pouquíssima evidência concreta, talvez nenhuma, em que fundamentar a reconstrução do segmento da oclusiva labial sonora *b* (observe a marca interrogativa na tabela (2.6) acima. Uma lacuna desse tipo é pouco comum nas línguas do mundo). (3) Nenhuma raiz indo-europeia existe na forma oclusiva sonora + vogal + oclusiva sonora: por exemplo, \**deg-* é um morfema impossível. Tal restrição na estrutura morfológica parece carecer por completo de motivação. Existem ainda outros problemas, mas os três apresentados acima serão suficientes para nossas finalidades atuais.

Como consequência dessas dificuldades, vários pesquisadores propuseram revisões da descrição fonética das oclusivas PIE. A maioria dessas revisões (vide, p. ex., Hopper, 1973) inclui uma reanálise das oclusivas sonoras tradicionais como oclusivas glotalizadas (ejetivas). Nas línguas que exibem uma série ejetiva, é comum o segmento labial ([p']) faltar. Além disso, as sequências do tipo ejetiva + vogal + ejetiva são raríssimas. Pois, a restrição contra os morfemas do tipo *deg-* não seria problemática se *deg-* fosse reinterpretado como *t'ek'-*. Finalmente, as oclusivas sonoras aspiradas (às vezes chamadas “oclusivas murmuradas”) ficam menos difíceis de integrar num sistema que não apresenta um contraste fonêmico de vozeamento. Já que as aspiradas sonoras são caracterizadas por um traço fonético designado “murmúrio” na análise fonética não tradicional, as três séries de oclusivas do PIE seriam distinguidas pelos traços de globalização e murmúrio. Compare as oclusivas em (2.6) com as do exemplo abaixo:

(2.7)

Surda/dezvozeada	Glotalizada	Murmurada (= sonora e aspirada)	
<i>p</i>	[ <i>p'</i> ]	<i>b</i>	(= <i>bh</i> )
<i>t</i>	<i>t'</i>	<i>d</i>	(= <i>dh</i> )
<i>k</i>	<i>k'</i>	<i>g</i>	(= <i>gh</i> )
<i>k<sup>w</sup></i>	<i>k<sup>w'</sup></i>	<i>g<sup>w</sup></i>	(= <i>g<sup>w</sup>h</i> )

A reconstrução apresentada acima em (2.7) ainda é controversa dentre os especialistas no indo-europeu, mas essa hipótese consegue demonstrar, de qualquer modo, a natureza de algumas das tentativas modernas de trazer um maior grau de exatidão para a reconstrução dos sistemas fonológicos.

## A hipótese da árvore genealógica

Como já tivemos ocasião de apontar, o método comparativo pressupõe a relação de parentesco entre as línguas a serem comparadas. Para pintar as qualidades especiais desse parentesco, August Schleicher introduziu o conceito da árvore genealógica em 1871. A noção revela o interesse da época na teoria de evolução e aplica a hipótese, no que diz respeito ao desenvolvimento de espécies diferentes, à evolução das línguas filhas a partir de uma língua ancestral. A hipótese da árvore genealógica supõe divisões sucessivas nas etapas anteriores razoavelmente homogêneas, períodos de desenvolvimento durante os quais outras mudanças possam ter ocorrido e ainda outras divisões. Pela ocorrência regular de tal série de eventos, as famílias linguísticas proliferam. Pressupõe-se que, após uma língua ancestral dividir-se em duas ou mais línguas filhas, os falantes das línguas filhas se dispersam por caminhos separados, linguística e, muitas vezes, fisicamente. Não há nenhum contato

mais entre os falantes das línguas filhas. De modo a exemplificar esse fenômeno, uma porção da árvore genealógica da família uraliana e uma versão reduzida da árvore genealógica da família IE são apresentadas nas figuras 2.1 e 2.2 abaixo.

Ao reconstruir a história de famílias linguísticas, é importante dispor de um método para estabelecer que uma língua se dividiu em duas ou mais línguas filhas. Em geral, a *cisão linguística* está baseada na noção da *reestruturação fonológica*. O sistema fonológico de uma língua é reestruturado quando o sistema de contrastes fonológicos for modificado de tal modo que antigos contrastes desapareçam, novas oposições sejam introduzidas, ou quando os elementos do sistema simplesmente sejam realinhados. Se um membro de um par de dialetos sofrer um desenvolvimento ou uma série de desenvolvimentos que reestruture seu sistema fonológico, é possível dizer que a língua original se dividiu em duas línguas. Em tal situação, a mudança é irreversível e um sistema fonológico é estabelecido, que é inovador, tanto com respeito ao sistema da língua ancestral, quanto é distinto dos sistemas das línguas cognatas.

No antigo índico, por exemplo, a duração era evidentemente distintiva nas vogais altas. Essa distinção desapareceu no bengalês, no entanto, como mostram os dados seguintes. O hindu representa a situação original, mantida numa língua índica moderna.

(2.8)

<b>bengalês</b>	<b>hindu</b>		
<i>ĵib</i>	जीभ ( <i>jībh</i> )		“língua”
<i>din</i>	दिन ( <i>din</i> )		“dia”
<i>dur</i>	दूर ( <i>dūr</i> )		“longe”
<i>ṣona</i>	सुन ( <i>sun-</i> )		“ouvir”

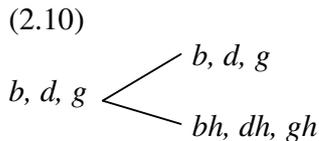
Podemos supor que o bengali e o hindu refletem uma cisão na história das línguas índicas à base dessa reestruturação (entre outras) no sistema fonológico do bengali.

Um desenvolvimento no sistema consonantal do irlandês antigo oferece mais um exemplo de natureza semelhante. As classes das oclusivas sonoras aspiradas e das oclusivas sonoras não aspiradas do protoindo-europeu se fusionam no celta primitivo, eliminando um contraste no irlandês antigo (irl. ant.) onde havia antes uma distinção no PIE e onde ainda existia uma diferença entre as duas classes de oclusivas em muitas das línguas irmãs indo-europeias do irlandês antigo, tal como o grego antigo, por exemplo, como fica evidente abaixo:

(2.9)

<b>PIE</b>	<b>irlandês antigo</b>	<b>grego antigo</b>	
* <i>g</i>	<i>gein</i> (/geni/)	γενος ( <i>genos</i> )	“nascimento”
* <i>g<sup>h</sup></i>	<i>géis</i> (/ge:ní/)	χην ( <i>khēn</i> )	“cisne”
* <i>d</i>	<i>deich</i> (/dexi/)	δεκα ( <i>deka</i> )	“dez”
* <i>d<sup>h</sup></i>	<i>dorus</i> (/dorus/)	θυρα ( <i>thyra</i> )	“porta”
* <i>b</i>	<i>toib</i> (/tobi/)	[latim: <i>tibia</i> . Não há cognato no grego.]	“osso”
* <i>b<sup>h</sup></i>	<i>berid</i>	φερω ( <i>pherō</i> )	“levar”

Tal como foi o caso no bengali, a inovação irlandesa é, de fato, uma instância de mudança irreversível. O falante do irlandês antigo não tinha como saber quais das oclusivas sonoras derivam das oclusivas sonoras aspiradas do PIE e quais são as descendentes das oclusivas sonoras não aspiradas. Além disso, já que não há nada na composição das palavras envolvidas que tenha podido indicar a situação anterior, até uma mudança hipotética dos tipos seguinte teria que se basear em fatores fonéticos ou ainda em outros indícios diferentes presentes no irlandês:

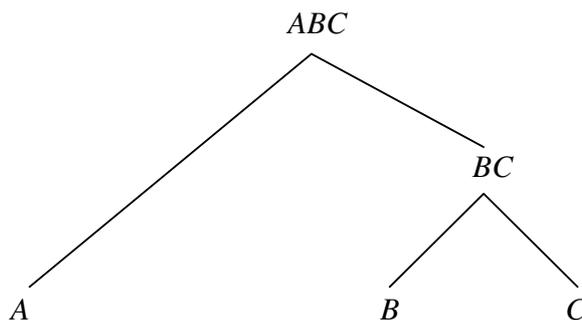


Pelo fato de um sistema fonológico recém-reconstruído se desenvolver de modo independente das suas origens, a instauração de uma cisão linguística é baseada em tal reestruturação.

Uma certa mudança fonológica pode ser refletida, naturalmente, em mais de uma língua aparentada. O fato de que várias línguas podem compartilhar determinadas mudanças implica a existência de subagrupações genéticas dentro de uma família linguística. Tal subagrupação é de se esperar se as famílias linguísticas espelham, de fato, uma série de cisões sucessivas, tal como Schleicher sugeriu. A relação genética entre algum subconjunto das línguas filhas de uma família maior é definida em termos da hipótese da árvore genética e é estabelecida por meio do método comparativo.

Ao conjecturarmos a existência três línguas filhas, *A*, *B* e *C*, podemos afirmar que as línguas *B* e *C* têm uma relação mais próxima do que qualquer uma dessas duas tem com a língua *A*, se um sistema fonológico puder ser reconstruído a partir do qual *B* e *C* podem ser derivadas, mas que é diferente do sistema fonológico do proto-*ABC*. Em tal caso, um nó intermediário pode ser estabelecido na árvore genealógica da seguinte maneira:

(2.11)



Os desenvolvimentos que diferenciam o sistema sincrônico do protoBC do sistema sincrônico do ABC são designados como as **inovações compartilhadas** (ou **comuns**), que caracterizam o período de desenvolvimento comum de BC.

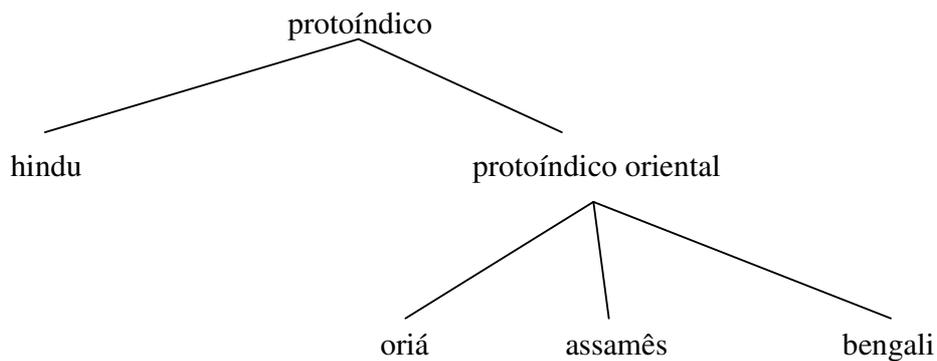
Em (2.12), cognatos de mais duas línguas foram adicionados aos dados fornecidos para o hindu e o bengali na apresentação anterior acerca da cisão linguística.

(2.12)

<i>oriá</i>	<i>assamês</i>	<i>bengali</i>	<i>hindu</i>
<i>ḷibhɔ</i>	<i>zibha</i>	<i>ḷib</i>	<i>ḷibh</i>
<i>dinɔ</i>	<i>din</i>	<i>din</i>	<i>din</i>
<i>durɔ</i>	<i>dur</i>	<i>dur</i>	<i>dūr</i>
<i>sun-</i>	<i>xun</i>	<i>šona</i>	<i>sun-</i>

Na base de todas as quatro línguas, reconstruiríamos quatro vogais altas, *i*, *ī*, *u*, *ū*. Entretanto, se desconsiderarmos o hindu, podemos reconstruir um sistema fonológico para as outras três línguas em que há somente duas vogais altas, *i* e *u*. Essas três línguas exibem uma inovação compartilhada na fusão das vogais altas longas e breves. Na base desse fenômeno (e de certas outras informações relevantes), essas três línguas são tidas como compondo um subgrupo na árvore genealógica índica, que divide um antecessor intermediário comum, o proto-oriá-assamês-bengali, também conhecido como o protoíndico oriental.

(2.13)



## A teoria das ondas

Existe um importante concorrente ao modelo do desenvolvimento linguístico da árvore genealógica que é chamado a **teoria das ondas**. Proposto primeiro por Johannes Schmidt no ano de 1872, a teoria das ondas sustenta que as inovações linguísticas se espalham de uma língua ou um dialeto para outra variedade linguística através dos contatos por parte dos falantes de línguas ou dialetos vizinhos, e os defensores dessa teoria demonstram que é frequente as línguas compartilharem inovações que não podem ser atribuídas a um antecessor comum. Duas línguas *A* e *B* poderiam exibir, por exemplo, reflexos de mudanças fonológicas idênticas. *A*, no entanto, pode exibir também os reflexos de uma outra mudança que são específicos à língua *A* e que têm que ter precedido o desenvolvimento que é idêntico ao da língua *B*. Uma mudança comum a mais de uma língua (sejam elas aparentadas ou não) que não é devida à herança de um antecessor comum é designada uma **evolução paralela**.

É possível demonstrar que muitas inovações conhecidas são o resultado de uma difusão ondulatória de uma língua ou de um dialeto para outro. A metafonía (*umlaut*) nas línguas germânicas, por exemplo, parece não refletir uma inovação protogermânica, mas parece ter se difundido pelos dialetos germânicos depois do período da diferenciação. A mais antiga língua germânica atestada, o

gótico, não exhibe a metafonía. O processo é encontrado tanto nas línguas germânicas setentrionais, quanto nas línguas germânicas ocidentais. Uma expansão em forma semelhante a uma onda é sugerida pelo fato de que os dialetos falados no sul da Alemanha exibem muitas formas sem metafonía em comparação com os dialetos falados no norte da Alemanha.

Embora a teoria da árvore genealógica e a teoria das ondas tenham sido tratadas por alguns linguistas como se fossem teorias em concorrência, não parece existir motivo para acreditar que se trate de hipóteses sobre a evolução linguística que se excluem mutuamente. Muitas mudanças podem ser reconstruídas para o protogermânico e essas modificações estão refletidas em todas as línguas germânicas – por exemplo, a fusão do *ā* e *ō* como *a* – mas outras mudanças, tal como a metafonía, se espalharam pela família. É evidente que essas duas hipóteses são complementares e que é necessário tomarmos o cuidado de diferenciar entre as mudanças que refletem a relação genética e as que refletem a difusão quando compararmos as línguas para fins de inferência histórica.

## A identificação de palavras emprestadas

O emprestar de vocábulos de uma língua para outra não deveria criar grandes dificuldades para a reconstrução comparativa se o item tomado emprestado foi introduzido de alguma língua totalmente não aparentada. Em tal caso, é provável que não haja cognatos nas línguas aparentadas e a forma tomada emprestada não aparecerá nos conjuntos correspondentes. No entanto, se uma palavra for tomada emprestada por uma língua de uma de suas parentes, é mais do que provável que existam cognatos e é esse tipo de situação de empréstimo que poderia provocar verdadeiros problemas para o pesquisador comparativista.

Já que as mudanças fonológicas individuais são restritas em geral a um determinado período definido na história da língua, não se pode esperar que um item exiba os efeitos de certa mudança fonológica se esse vocábulo foi emprestado à língua receptora depois de tal mudança fonológica ter acabado de se operar. É por força desse fenômeno que o linguista pode identificar quais itens foram tomados emprestados. Se uma forma não se conformar (em todos seus detalhes) ao esquema de correspondências fonológicas que tenha sido estabelecido na base de números significativos de itens, tal forma pode ser tomada por suspeita. Considere o exemplo seguinte em que uma palavra foi tomada emprestada de uma língua irmã.

A forma eslava *gōsĩ* (> *гусь* (*gus'*) “ganso” em russo) é evidentemente uma palavra indo-europeia. No entanto, a oclusiva velar sonora inicial (*/g-/*) não é de antecipar numa língua eslava. Existem cognatos em latim (*anser* < *hanser*), em sânscrito ( **हंस** (*hansa*)), em grego (**χην** (*khēn*)) e no germânico (*\*gans*, o antecessor reconstruído de formas como *goose* (ingl.), *Gans* (alem.), *gás* (norreno)). Na base dos cognatos latino, sânscrito, grego e germânico, é possível reconstruir *\*ǵ<sup>h</sup>ans* no protoindo-europeu. Os segmentos *g* do germânico, *h* do latim, *h* do sânscrito e *k<sup>h</sup>* do grego se correspondem regularmente em posição inicial, todos sendo deriváveis do segmento *\*ǵ<sup>h</sup>* do PIE. O *g* eslavo atestado em *gōsĩ*, porém, não reflete *\*ǵ<sup>h</sup>* PIE. O reflexo antecipado no ramo eslavônico da família indo-europeia é *z*, tal como pode ser observado nos conjuntos cognáticos como, por exemplo, *YIΓAH* (*-wigan*) do gótico, *VEHO* (*vehō*) em latim, **वह-** (*vah-*) do sânscrito, **ὄχος** (*okhos*) do grego e **βεζο** (*vezō*) do eslavônico, que significam “trazer”; ou *HIEMS* (lat.), **हेमन्त-** (*hemanta-*) (sânc.), **χείμα** (*kheima*) (gr.) e **зима** (*zima*) (esl.) para “inverno”, ou *HUMUS* (lat.), **χάμαι-** (*khamai-*) (gr.) e **земля** (*zempl'a*) (esl.) para “terra”. Em virtude dessas correspondências cognáticas, portanto, *gōsĩ*

não pode representar o desenvolvimento regular de \*ǵ<sup>h</sup>ans do PIE no ramo eslavo e, geralmente, supõe-se que o termo eslavônico reflete um empréstimo de um dos dialetos germânicos vizinhos.

No caso de *gōsī* no eslavo, a palavra aberrante tinha sido tomada emprestada de uma língua irmã (o germânico). Também é possível que uma palavra seja introduzida, ou, melhor, reintroduzida, a uma língua filha a partir da sua língua de origem ou de algum dialeto estreitamente aparentado à língua receptora. Tal situação se desenvolve com frequência sob a influência de uma tradição literária vigorosa. Dois exemplos do hindu moderno exemplificam essa situação.

Em algum momento entre o período do índico antigo (sânscrito) e o surgimento do hindu moderno, as obstruintes simples intervocálicas ou foram eliminadas, ou se converteram em glides, p. ex., भ्रातर (bhrātar) (sânsc.) “irmão” > भाई (bhāī /b<sup>h</sup>a:ɪ/) (hindu) “irmão”; चतुर्थ (caturtha) (sânsc.) “quatro” > चउथ (cautha /cauth<sup>h</sup>a/) (hindu) “quatro”. Uma obstruinte intervocálica é mantida, porém, na palavra hindu राजा (rājā) “rei”. Essa palavra é frequente no antecessor do hindu, o sânscrito, e diversos cognatos existem em outras línguas indo-europeias, por exemplo, REX (rēx /re:ks/) em latim. Sua ocorrência no hindu pode ser explicada somente se for tratado como uma reintrodução por meio da influência do sânscrito literário arcaico, num período depois que a perda das obstruintes intervocálicas parou de operar.

A palavra hindu नाम (nām) é igualmente afastada dos padrões da evolução fonológica. A forma नामन- (nāman-) ocorre no sânscrito, tornando-se em नाम (nām) durante o período do índico médio. O hindu exibe um processo desenvolvimental regular de /-Vm/ > [-Ṽw̃] / \_\_ #, por exemplo, ग्राम (grāma) “aldeia” (sânsc.) > गाम (gām) (índ. méd.) > गाउं (gāw̃). A maioria dos dialetos estreitamente aparentados ao hindu apresentam tanto गाउं (gāw̃), como नाउं (nāw̃). Desse modo, a presença de नाम (nām) no hindu deve representar uma outra situação, em que a língua contemporânea aceitou um empréstimo de uma língua que representa uma fase anterior da sua própria história.

## Leituras adicionais recomendadas

BLOOMFIELD, L. (1922). *Language*. New York: Holt, Rinehart and Winston, chapter 18.

HAAS, M. R. (1966). “Historical Linguistics and Genetic Relationship”, in: *Current Trends in*

*Linguistics*, vol. 3: 113-53. The Hague: Mouton.

HALL, R. A. (1950). “The Reconstruction of Proto-Romance”, *Language* 26: 6-27.

HOCKETT, C. F. (1948). “Implications of Bloomfield’s Algonquian Studies”, *Language* 24: 117-31.

HOENIGSWALD, H. (1963). “Criteria for Sub-grouping Languages”, in: H. BIRNBAUM and J.

PUHVEL (eds), *Ancient Indo-European Dialects*. Los Angeles, CA: University of California Press.

Curso de Linguística Histórica (FLL 0443) – Prof. Thomas Finbow – Departamento de Linguística  
Faculdade de Filosofia, Letras e Ciências Humanas – Universidade de São Paulo (FFLCH/USP)

HOENIGSWALD, H. (1950). “The Principal Step in Comparative Grammar”, *Language* 26: 357-64.

HOPPER, P.J. (1973). “Glottalized and Murmured Oclusives in Indo-European”, *Glossa* 7.2: 141-66.

MEILLET, A. (1970). *Le méthode comparative en lingistique historique*. Paris: H. Champion.