

## A TIPOLOGIA LINGUÍSTICA

A tipologia linguística estuda as línguas em um modo transversal e as classifica não como derivadas de um ancestral comum (como na abordagem histórico-comparativa), mas baseado na pertinência a um tipo estrutural comum.

A tipologia se propõe de entender até que ponto as línguas podem ser e são diferentes, quanto às estratégias usadas para expressar significados. Ela estuda a estrutura das línguas independentemente de sua dimensão histórica. Desse ponto de vista, todas as línguas têm o mesmo interesse, sendo essas vivas, mortas, atestadas ou não. O hitita, portanto, tem a mesma legitimidade do inglês aos fins da tipologia. Os grupos entre os quais as línguas são divididas se chamam tipos linguísticos. O **tipo linguístico** é uma combinação de traços estruturais logicamente independentes. São consideradas pertinentes aquelas propriedades que permitem fazer previsões atendíveis sobre a estrutura das línguas pesquisadas. Isto é, se um tipo é um conjunto de propriedades independentes, mas relacionadas, qualquer uma dessas propriedades será pertinente se permitir a previsão das outras.

Façamos um exemplo tirado de outro âmbito. Imaginemos ter que classificar o conjunto dos veículos a motor de modo a poder fazer previsões sobre suas características e formas. Pode-se partir do destino de uso, dividindo os veículos em três tipos: transporte de passageiros, traslado de animais e carregamento de mercadoria. Os veículos dos últimos dois tipos provavelmente serão bastante grandes e terão dispositivos para o reboque da carga ou, no caso dos animais, uma jaula separada da cabine de direção. Os do primeiro tipo teriam, entre características essenciais, a presença de assentos, a ausência de uma cabine de direção separada do resto do veículo, etc. Em suma, podemos dizer que o parâmetro do destino de uso permite fazer previsões sobre outros parâmetros. Se, ao contrário, tivéssemos escolhido o parâmetro cor, não teríamos conseguido deduzir outras características, porque qualquer tipo de veículo pode ser de qualquer cor.

Um parâmetro linguístico que demonstrou grande capacidade de previsão é aquele relativo a algumas estruturas sintáticas, como o SV (sintagma verbal), SN (sintagma nominal), a frase relativa e outros. Partindo da ordem dos constituintes de uma dessas estruturas, pode-se prever com boa probabilidade a ordem dos constituintes das outras.

A tipologia não é uma disciplina puramente descritiva, mas preditiva. O tipo tem caráter meramente estrutural e a tipologia deve poder individualizar o princípio subjacente que relaciona as várias propriedades correlatas. Não são relevantes as características isoladas, mas a *ratio* profunda, o princípio mais profundo que explica as relações que intercorrem entre as várias características. Isso não quer dizer que as características apareçam sempre e com certeza, mas que é provável que apareçam. De fato as línguas estão sujeitas, no seu devir histórico, a muitos condicionamentos que vão além da simples realidade linguística e que podem imprimir bruscos e imprevisíveis desvios. Os tipos são simplificações da multiplicidade que existe na natureza e, portanto, não reproduções integrais e fiéis de nenhuma língua. Uma língua será inscrita em um tipo ou em outro se a

sua estrutura for estatisticamente compatível mais a um tipo que a outro. Obviamente não devem existir na realidade tipos que teoricamente resultem impossíveis.

Vejamos qual é a explicação mais plausível das causas subjacentes às correlações estruturais de cada tipo. A causa seria externa ao sistema linguístico e seria de natureza pragmática: visto que a língua é, em última análise, um fato social e que a sua principal função é possibilitar a comunicação nas sociedades, a *ratio* profunda seria de natureza funcional. Cada configuração tipológica seria então um reflexo das estratégias de que a língua predispõe para resolver os problemas comunicativos.

Nos estudos tipológicos há uma clara prevalência dos estudos que se ocupam da morfologia e da sintaxe. O léxico é de fato uma espécie de revestimento externo da língua, muito em contato com a mutabilidade da realidade extralinguística para poder fornecer indicações confiáveis à reconstituição dos tipos. O nível fonético-fonológico é, ao contrário, o nível da língua mais impermeável às influências externas e, por isso mesmo, ficou descuidado, porque se manteve muito conservador.

### **Formas marcadas *versus* não marcadas**

Na língua, em todos os níveis descritivos, temos formas chamadas de mais marcadas e formas chamadas de menos marcadas. Esse conceito de marcado e não marcado pode ser interpretado de duas maneiras, mas na maior parte das vezes as duas maneiras se somam. A primeira maneira é interpretar marcado com mais raro estatisticamente. Portanto o conceito de marcado seria ligado a uma menor frequência. A segunda maneira é interpretar marcado, como diz a própria palavra, como algo que carrega mais marcas gramaticais do que seu equivalente não marcado. Veremos que, normalmente, o que acontece é que temos os dois conceitos juntos. Além disso, é importante saber que a existência de uma forma marcada implica a existência de um equivalente não marcado e que se aprende antes a forma não marcada e depois a marcada.

Olhando na fonética, podemos dizer que as vogais nasais do português são marcadas, ou seja, são uma forma mais rara e/ou com mais traços gramaticais do que as vogais orais. As vogais *ã*, *õ* etc. são mais marcadas do que *a*, *o* etc. Do ponto de vista da frequência, as vogais nasais são menos frequentes do que as orais. Todas as línguas têm vogais orais, mas somente algumas têm vogais nasais. Mesmo nas línguas que têm vogais nasais, as orais são mais frequentes. As vogais nasais possuem mais traços do que as orais, pois aos traços presentes nas correspondentes orais precisa acrescentar o traço nasal. As vogais nasais são aprendidas depois das vogais orais, e normalmente são um elemento de difícil aprendizado para os estrangeiros. Uma das principais dificuldades que os estrangeiros têm em aprender o português está nas vogais nasais.

Outro exemplo de traço fonético marcado é a oposição de quantidade consonantal, típica do italiano. A oposição fonológica entre consoantes longas e breves é um traço raro nas línguas e muito difícil de ser adquirido pelos estrangeiros. Mesmo nas línguas que possuem esse traço, não se verifica em todas as consoantes nem em todas as posições. Essa

oposição acrescenta um traço a mais, pois não se trata somente de distinguir entre uma consoante e outra, mas também entre a quantidade delas.

Outro exemplo de traço fonético marcado é a presença de consoantes vozeadas em final de palavra. O inglês é uma língua que possui esse traço. Todas as línguas que possuem consoantes vozeadas em final de palavra possuem também, na mesma posição, as desvozeadas. Ao contrário, há línguas, como o alemão, que possui as finais desvozeadas, mas não as vozeadas.

O conceito de marcado e não marcado é aplicável também aos outros níveis descritivos. Por exemplo, podemos dizer que a forma não marcada de plural em português é com -S, mas existem também outras formas, mais raras e com mais traços. Podemos dizer que a categoria de número é menos marcada do que a categoria de gênero. Se uma língua possui a categoria de gênero, inevitavelmente possui aquela de número.

Na sintaxe podemos notar, por exemplo, que certas construções são mais frequentes e menos marcadas que outras. Um exemplo clássico é o caso das relativas. Não todas as línguas possuem as construções relativas, mas se uma língua possui um tipo de construção relativa é a relativa do sujeito (por exemplo: *A garota que fala é minha irmã*); se as línguas possuem dois tipos de relativas possuem aquela do sujeito e aquela do objeto direto (*A garota que você cumprimentou é minha irmã*); se possuem três, a terceira é do dativo (*Dou um presente ao amigo que eu amo*); se possuem quatro, a quarta é do genitivo (*Compro a casa cujo preço é melhor*). Em prática, a última implica a presença das anteriores. Assim a do sujeito é a menos marcada e a do genitivo a mais marcada.

## A tipologia sintática

Vejam os elementos da frase declarativa são dispostos. Os constituintes básicos são S(ujeito), V(erbo) e O(bjeto). Emerge uma clara prevalência de dois tipos: 45% das línguas são SOV (Japonês, Turco, parte das línguas ugro-fínicas, línguas dravídicas, etc.) e 42% são SVO (línguas românicas, germânicas, eslavas, línguas bantas, vietnamita, etc.). Menos de 10% adota a ordem VSO.

Japonês SOV

*Taro ga inu o mita*

Taro cão viu

Taro viu o cão

Yoruba SVO

*Bàbá ra bàtá*

Pai comprou sapatos

O pai comprou os sapatos

Gallese VSO

*Lladdodd y ddraig y dyn*

Matou o dragão o homem

O dragão matou o homem

Estes três tipos cobrem cerca de 97% da variação interlinguística mundial. O que aconteceu aos outros três tipos (VSO, OSV e OVS)? São VOS o malgaxe (Madagascar) e o coeur d'Alène (língua penuti falada ao norte de Salt Lake City), enquanto é OVS o hixkaryana (Brasil). Ao contrário dos outros tipos, é possível que não existam línguas OSV.

A preferência das línguas do mundo pelo tipo SVO e SOV (e um pouco pelo tipo VSO) é tão clara que não pode ser fruto do acaso. O elemento que os três tipos preferidos têm em comum é a ordem SO. Então, podemos dizer que (quase) todas as línguas do mundo colocam o S antes do O na frase principal declarativa. Qual pode ser a razão? O S, geralmente, é a entidade que realiza a ação expressa pelo verbo e que exerce sobre essa um alto grau de controle. Estas prerrogativas conferem ao S uma saliência cognitiva superior ao O na organização mental da informação. Notamos, de fato, que a primeira posição de uma frase confere ao constituinte uma relevância maior no ato comunicativo. Além disso, o S corresponde na maior parte dos casos àquela parte da informação que é definida como dada, enquanto, no resto da frase, se transmite a informação nova.

### **Ordem natural e ordem marcada**

Na realidade a situação é mais complicada. Algumas línguas possuem uma ordem dos constituintes muito rígida (o inglês) e outras têm uma ordem muito livre (o húngaro). Relativamente livre é a ordem do Italiano:

*Carlo prese la giacca*

Carlos pegou o casaco

*La giacca, la prese Carlo*

O casaco, o pegou Carlos

*La prese Carlo, la giacca*

O pegou Carlos, o casaco

*Prese la giacca, Carlo*

Pegou o casaco, Carlos

O que diferencia a frase *Carlos pegou o casaco*, que está na ordem SVO, de todas as outras? A primeira é natural, não marcada, isto é, serve para todos os contextos, tem mais probabilidade de ser usada e não requer um contexto particular. As outras têm menos

probabilidade de serem usadas e requerem um contexto preciso, ou um número mais reduzido de contextos, além de uma ou mais marcas gramaticais a mais. Por exemplo, a frase não marcada pode responder a todas as seguintes perguntas:

Quem pegou o casaco?

O que Carlos pegou?

O que fez Carlos?

O que aconteceu?

*Carlo pegou o casaco*

Ao contrário, à pergunta *Quem pegou o casaco?* não posso responder *\*Pegou o casaco, Carlo* e devo responder *O casaco, o pegou Carlos*. E à pergunta *O que fez Carlos?* não posso responder *\*O casaco, o pegou Carlos*.

Quando uma forma é marcada, quer dizer que exprime um valor informativo não natural, mas específico de um contexto, que é menos provável e que pressupõe a existência da forma não marcada (enquanto a forma não marcada não pressupõe a existência da marcada) e que traz consigo uma marca gramatical especial.

Na forma natural, a frase é pronunciada sem nenhuma marca especial. Nas outras formas, ao contrário, há uma quebra e uma entonação particular que é assinalada com a vírgula. É justamente esta a marca gramatical particular.

Há muitos parâmetros correlatos à ordem VO e também à OV. Isto é, sabendo se uma língua é OV ou VO, pode se prever muitas outras características. Por isso, o parâmetro da ordem dos constituintes básicos é muito potente.

**TIPO VO:** N-art., preposições, N-genitivo, N-Adj., N-demonstrativo, N-numeral, N-FRel., Aux-V, V-Adv., Conj.-Subordinada, Comparativo-Compl. de comparação, Pron. de interrogação em posição inicial.

**TIPO OV:** art.-N, posposições, genitivo-N, Adj-N, demonstrativo-N, numeral-N, FRel.-N, V-Aux., Adv.-V, Subordinada-Conj., Compl. de comparação-Comparativo- Pron. de interrogação no final.

**TIPO VO (Tailandes):**

*Khǎw sii aahǎan*

Ele compra comida

Preposições

*Kàp chaawbâan*

Com pessoas da mesma aldeia

N-genitivo

*Nānsīi dèk*

Livro menino (o livro do menino)

N-Adj.

*Bāan sūay*

Casa bonita (uma casa bonita)

N-Demonstrativo

*Wichaa nān*

Argumento este (este argumento)

N-Numeral

*Dèk sām khon*

Menino três CLASSIFICADOR (três meninos)

N-FRel.

*Dèk thīi rian phaasāa thay maa léew*

Menino REL estuda língua tailandesa chega já

O menino que estuda a língua tailandesa já chegou

Pron. de interrogação inicial

*Khray rian phaasāa thay?*

Quem estuda língua tailandesa?

Quem estuda a língua tailandesa?

**TIPO OV (Turco):**

*Hasan ökuzu aldi*

Hasan o boi comprou (Hasan comprou o boi)

Posposições

*Masanin altinda*

A mesa embaixo (embaixo da mesa)

Genitivo- N

*Ayşe-nin arabasi*

Ayşe-GEN carro (o carro de Ayşe)

Adj.- N

*Yorgun at*

Cansado cavalo (um cavalo cansado)

Demonstrativo- N

*Bu makale*

Este artigo

Numeral- N

*Beş adam*

Cinco homem (cinco homens)

FRel.- N

*Okula gid-en çocuk*

Escola vai-REL menino (o menino que vai à escola)

Pron. de interrogação final

*Çokuğa kitabı kim verdi?*

Ao menino o livro quem deu? (Quem deu o livro ao menino?)

Os dois tipos em questão indicam o índice máximo de coerência respectivamente em berbere e zapoteco para as línguas VO e em basco, birmano e hindi para o tipo OV. No entanto, esses não cobrem toda a gama de variedade interlinguística.

Por simplicidade argumentativa, vejamos somente os quatro primeiros parâmetros da lista: ordem de V e O, prep./posp., posição de Adj e N, posição de Gen. e N.

Somente sobre esses quatro parâmetros os tipos possíveis são quinze, mas apenas sete são realmente significativos estatisticamente. E ainda, os tipos logicamente possíveis seriam várias dezenas. Isto não pode ser uma casualidade.

Vejamos o português e o inglês (duas línguas geneticamente muito próximas):

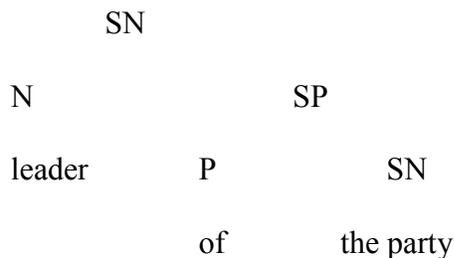
<b>PORTUGUÊS</b>	<b>INGLÊS</b>
<b>VO</b> Ele compra comida	<b>VO</b> He buys food
<b>Preposições</b> <i>Para o Brasil</i>	<b>Preposições</b> <i>To Brazil</i>
<b>N-genitivo</b> <i>O livro de Pedro</i> <i>O tempo da nossa vida</i>	<b>N-genitivo</b> <i>The time of our life</i> <b>genitivo-N</b> <i>Pedro's book</i>
<b>N-Adj.</b> <i>Um homem pobre</i> <b>Adj.-N</b> <i>Um pobre homem</i>	<b>Adj-N</b> <i>A poor man</i>
<b>Demonstrativo-N</b> <i>Este homem</i>	<b>Demonstrativo-N</b> <i>This man</i>
<b>Numeral-N</b> <i>Três homens</i>	<b>Numeral-N</b> <i>Three men</i>
<b>N-FRel.</b> <i>O menino que estuda historia</i>	<b>N-FRel.</b> <i>The man who studied history</i>
<b>Pron. de interrogação inicial</b> <i>Quem estudava história?</i>	<b>Pron. de interrogação inicial</b> <i>Who studied history?</i>

Trata-se de duas línguas SVO. O português não respeita os parâmetros previsíveis em dois casos (N-Demonstrativo e N-Numeral) e em um caso (N-Adjetivo) pode ter as duas ordens. O inglês não respeita os parâmetros em três casos (N-Adjetivo, N-Demonstrativo e N-Numeral) e em um caso (N-genitivo) pode ter as duas ordens. Se acrescentamos a estas duas línguas o italiano, percebemos que a ordem ideal não é respeitada no N-Demonstrativo e N-Numeral e que para N-Adjetivo tem-se duas ordens. Então, temos em três línguas sobre três alguns parâmetros que parecem particularmente inclinados a trair o tipo: N-Demonstrativo, N-Numeral e, em parte, N-Adjetivo.

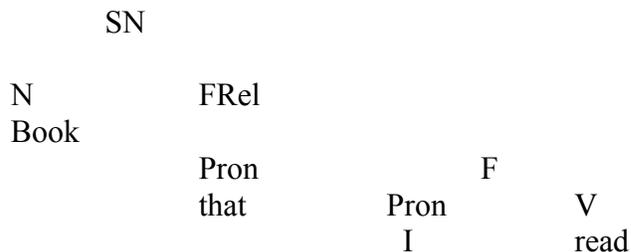
Decomponhamos parte por parte o primeiro tipo, o VO. No que concerne ao SV, o elemento dominante, tecnicamente definido núcleo, é o verbo. Tanto é verdade, que este

decide se deve haver ao menos um objeto direto e como este deve ser: podemos dizer *o menino come a maçã*, ou ainda *o menino come*, ou *come a maçã* mas não *\*o menino a maçã*. Neste tipo o V precede o O, isto é, o seu complemento; mesmo no SN o núcleo, isto é, o N (que governa o acordo e que não pode ser omitido, como podem ser os seus modificadores), precede os próprios complementos e os próprios modificadores; no S adposicional, a adposição (isto é, o núcleo, mesmo que nem todos estejam de acordo que uma adposição possa ser considerada núcleo) precede o complemento. No tipo OV acontece o oposto. Então o princípio organizador dos dois tipos concerne à recíproca posição do núcleo e complementos/modificadores. O tipo VO obedece ao princípio “núcleo à esquerda” e o tipo OV ao “núcleo à direita”. Isto tem consequências importantes, seja do ponto de vista teórico ou empírico. Do ponto de vista teórico é significativo que algumas construções logicamente independentes obedeçam a um princípio organizador único e coerente. Do ponto de vista empírico, o critério permite aos falantes uma notável economia de energia em fase de aquisição e de uso. No entanto, esta explicação não explica porque o adjetivo e o artigo não são colocados no princípio. O artigo, como modificador, deveria vir antes no tipo VO e depois o N no tipo OV; o contrário acontece com frequência. Uma tentativa de explicar o problema prevê que a coerência tipológica seria rigorosamente respeitada somente pelos constituintes que têm uma estrutura (micro) sintática interna, isto é, pelos constituintes que em uma representação na árvore exibissem uma ramificação. Vejamos o exemplo do Inglês:

a) leader of the party “líder do partido”



b) book that I read “livro que eu li”



c) black dog “cachorro preto”

	SN	
Adj		N
black		dog

d) the/a table “a/uma mesa”

	SN	
Art		N
The/a		table

Enquanto no caso de N e Gen e de FRel há ramificação, no caso de Adj e de Art não há ramificação e, então, não há obrigação de respeitar a ordem.

### Tipologia e morfologia

A tipologia morfológica pressupõe a ação de dois parâmetros: o índice de síntese e o índice de fusão. O primeiro diz respeito ao número de morfemas indivisíveis no interior de uma palavra; o segundo diz respeito à segmentabilidade da própria palavra, isto é, o grau de dificuldade com o qual são individuadas as fronteiras entre os morfemas e, então, entre as menores unidades de significado das quais são portadores.

A combinação dos dois parâmetros consente dividir as línguas em pelo menos quatro tipos: isolante, polissintético, aglutinante e fusivo (flexivo).

Nas línguas **isolantes** o índice de síntese tem o valor mínimo: cada palavra tende a ser monomorfemática, cada morfema é invariável na forma e em geral exprime um único significado. Não tem sentido, neste caso, falar de índice de fusão, porque há um único morfema. Entre as línguas isolantes estão o mandarim e o vietnamita.

Chinês mandarim

*Tā zài túshūguǎn kàn bào*

Ele na biblioteca ler jornal

Ele está lendo um jornal na biblioteca

Vietnamita

*Khi tôi đến nhà bạn tôi, chúng tôi bắt đầu làm bài*

Quando eu ir casa amigo eu PL eu tomar cabeça fazer aula

Quando foi à casa do meu amigo, começamos (tomar cabeça=começar) a fazer a aula

Nestas frases cada palavra é formada de um único morfema. Por exemplo, observe-se o morfema *tôi* que vale “eu” no primeiro caso e no segundo vem pluralizado com o acréscimo do morfema *chúng* que traz apenas informação gramatical [PLURAL]. Ao contrário, no inglês o plural é assinalado pelo sufixo –s: *boy* “rapaz” e *boys* “rapazes”.

Além disso, as palavras tendem a ser variáveis na forma, contrariamente ao que acontece no português, em que a categoria sintática de uma palavra pode ser modificada por um prefixo e/ou sufixo (Adj. *Velho* - N *velhice* - V *envelhecer*). Nas línguas isolantes uma mesma palavra pode desempenhar diversas funções sintáticas, permanecendo formalmente tudo igual. Nas três ocorrências da forma *tôi* na frase vietnamita, na primeira e na terceira vez, essa é pronome pessoal, enquanto na segunda é adjetivo possessivo. Fenômenos deste tipo se chamam “conversão”, isto é, mudança de função morfossintática sem modificação formal da palavra. Existem também em línguas não isolantes: em português *rápido* pode ser usado como adjetivo (essa pessoa é rápida) e como advérbio (faça isso rápido). O processo da conversão é uma das características do tipo isolante. Exemplos italianos de conversão são a passagem de V a N (*cantante* particípio > *il cantante* N; *potere* “poder” infinitivo > *il potere* N “o poder”; *veduta* “vista” particípio > *la veduta* N; *crescendo* gerúndio > *il crescendo* N), ou a passagem de adjetivo a advérbio (*colore chiaro* > *parlar chiaro* “cor clara” > “falar claramente”). Mas é o inglês a língua não isolante que mais usa a conversão: *stop* pode ser N e V; *light* pode ser Adj., N, V, Adv.

Outro aspecto fundamental das línguas isolantes é a correspondência biunívoca entre os morfemas (neste caso palavras) e as unidades semânticas: cada morfema/palavra exprime um único significado, seja ele lexical ou gramatical. Por exemplo, a combinação dos dois significados [pronome pessoal de 1ª pessoa] e [plural], que em português está contida em uma única palavra (*nós* vs. singular *eu*) e o inglês exprime com *we* vs. singular *I*, em vietnamita é expressa colocando lado a lado dois morfemas/palavras: *chúng* (plural) e *tôi* (eu).

Em oposição às línguas isolantes encontramos as línguas **polissintéticas**, nas quais o índice de síntese assume o valor máximo. Essas, na verdade, concentram no interior da própria unidade lexical um grande número de morfemas, chegando a condensar em uma única palavra as informações que para nós requereriam uma frase inteira.

### **Esquimó**

*Angya-ghlla-ng-yug-tuq*

Barco-AUMENTATIVO-comprar-DESIDERATIVO-3ªPS. SING

Ele quer comprar um barco grande.

A sequência *Angyaghllangyugtuq* é ao mesmo tempo uma palavra e uma frase. Nestas línguas o índice de fusão tem valor intermediário porque o grau de complexidade da

estrutura interna das palavras faz com que haja casos em que morfemas adjacentes se fundem uns com os outros.

As línguas do tipo **aglutinante** têm um menor índice de fusão. A palavra consta de mais morfemas e a segmentação não apresenta dificuldades particulares. É rigorosamente mantida uma correspondência biunívoca entre forma e conteúdo.

**Turco:**

*Adam* (homem)

*Adam-lar* (homem-PL = homens)

*Adam-a* (homem-DAT = ao homem)

*Adam-lar-a* (homem-PL-DAT = aos homens)

**Nahuatl:**

*No-kali* (meu-casa = a minha casa)

*No-kali-mes* (meu-casa-PL = as minhas casas)

*Mo-pelo* (teu-cachorro = o teu cachorro)

*Mo-pelo-mes* (teu-cachorro-PL = os teus cachorros)

O índice de síntese nas línguas aglutinantes é médio-alto: as palavras tendem a ter muitos morfemas porque não é possível exprimir mais de uma categoria semântico-funcional em um único morfema.

Oposta é a situação das línguas flexivas ou **fusivas**. Aqui o valor do índice de fusão é máximo. Os limites entre um morfema e outro perdem em visibilidade e isto determina uma série de reações em cadeia: a segmentação se torna particularmente difícil, as exceções, digamos assim, se multiplicam e o ideal de correspondência biunívoca entre forma e conteúdo se perde porque mais categorias semântico-funcionais fundem-se em um único morfema. As línguas indo-europeias têm um caráter prevalentemente fusivo.

Se se considera o caso do dativo plural do latim *hominibus*, podemos segmentar inicialmente *homin-ibus*, com um morfema lexical e um morfema gramatical. Mas não é possível fazer outras segmentações. O morfema *-ibus* traz tanto o significado de dativo quanto de plural, portanto 2 significados correspondem à uma única forma. A diferença em relação ao turco é evidente: no dativo plural *adamlara*, *-lar* é a marca do plural e *-a* é a do dativo.

Nas demais línguas fusivas como o latim, os morfemas têm muitas vezes uma forma variável: a terminação do nominativo plural *-es* em *homines*, que pertence à 3ª declinação, não pode ser a mesma para os nomes masculinos de 2ª declinação, para os quais a mesma função é desempenhada pelo morfema *-i* de *lupi*, nem mesmo para os nomes neutros da 3ª declinação, para os quais o nominativo plural é feito com *-a* (*corpora*) ou para os nomes femininos da 1ª declinação (*rosae*). A violação da correspondência entre forma e conteúdo acontece em ambas as direções: por um lado *-ibus* é uma forma para dois significados

(plural e dativo), por outro *-ibus* pode ser interpretada tanto como dativo quanto como ablativo.

Nas línguas fusivas o índice de síntese se caracteriza por valores médio-baixos. A possibilidade de fazer convergir mais unidades semânticas em um único morfema consiste em reduzir o número de morfemas no interior da palavra.

Costuma-se encontrar no interior do tipo fusivo o subtipo introflexivo. Neste subtipo a relação entre unidade de forma e unidade de conteúdo é como o das línguas fusivas, porém as marcas morfológicas não vêm dispostas linearmente, mas em ordem de pente. O árabe, por exemplo, constrói as palavras intercalando uma raiz (tri)consonântica, que tem uma leitura semântica frequentemente genérica, desprovida de qualquer implicação gramatical (tempo, gênero, número, etc.) e sequências vocálicas particulares que são colocadas entre as consoantes, das quais se esperam as seguintes especificações:

*Ktb* = a área semântica da escrita

*Kataba* = ele escreveu

*Yaktubu* = ele escreve

*Kitāb* = livro

*Kutub* = livros

*Kutayyib* = livrinho

*Kātib* = escritor

*Kutubī* = vendedor de livros

*Maktaba* = biblioteca

*Miktāb* = máquina de escrever

## **Tipologia e fonologia**

Esse componente gramatical foi por muito tempo penalizado na análise tipológica, mas com algumas exceções. Uma delas é relativa a um elemento suprasegmental, o tom. O tom é uma propriedade que caracteriza os sons vozeados. Quanto maior a frequência da vibração das pregas vocais tanto mais alto o tom do som produzido. O tom se realiza em todas as línguas, mas com modalidades diferentes. Em cerca da metade das línguas o tom é fonologicamente pertinente, ou seja, tem valor distintivo. O tom permite que as palavras mudem de sentido. No chinês mandarim tem 4 tons, além do tom neutro. Na representação do tom se recorre a valores numéricos. Convencionalmente se supõe que a extensão máxima das variações de tom seja incluída num intervalo numérico de 1 (o ponto mais baixo) a 5 (o ponto mais alto). Os quatro tons do chinês mandarim são representados assim:

Tom 1: alto constante 55 -  
 Tom 2: alto ascendente 35 ´  
 Tom 3: descednente-ascendente 214 ˇ  
 Tom 4: alto descendente 51 `

As palavras seguintes se diferenciam unicamente pelo tom:

Palavra	IPA	
yī	[i] <sup>55</sup>	“vestido”
yí	[i] <sup>35</sup>	“suspeitar”
yǐ	[i] <sup>214</sup>	“cadeira”
yì	[i] <sup>51</sup>	“significado”

As línguas tonais são concentradas na África subsaariana, na América central e na Ásia sul-oriental. Portanto, fonologicamente podemos fazer uma distinção entre dois macrotipos: as línguas tonais e as não tonais. Internamente o tipo tonal não é homogêneo e pode ser reorganizado com base em diferentes parâmetros. Os principais dizem respeito à unidade à qual é associado o tom e a função que o tom tem. A distinção mais frequente é aquela que associa o tom a uma vogal (somalo *qaálin* “jovem camelo” vs *qaalín* “camela”) ou a uma sílaba (como nos exemplos do chinês). Dentro das línguas que associam o tom a uma sílaba é possível distinguir entre tons associados a uma única sílaba e tons que podem cobrir mais de uma sílaba (em digo, uma língua banto da Tanzânia, *a na ramuka* “(ela) está acordando” tem um tom alto distintivo nas últimas duas sílabas).

Quanto à função dos tons, a subdivisão mais importante é entre tons que distinguem morfemas lexicais e aqueles que distinguem morfemas gramaticais. Em aghem, uma língua falada nos Camarões, a preposição *à* “com” se diferencia unicamente pelo tom da preposição *â* “para”. Tem línguas em que o tom diferencia os números ou os tempos verbais. Os tons podem também ter uma função derivacional: em lendu, falado entre Congo e Uganda, *dhù* “insultar” e *dhú* “insulto”.

### Tipologia lexical

O léxico joga um papel decisivo na genética linguística, mas é marginal na tipologia. Enquanto as unidades fonológicas parecem refratárias a condicionamentos extra-sistêmicos, o léxico é absolutamente vulnerável ao contato. Mas mesmo assim existem setores semânticos mais impermeáveis e conservadores. Um caso é o da terminologia para as cores. Foram individuadas 11 cores que parecem ser reconhecidas e indicadas da mesma maneira pelos falantes das mais de 100 línguas escolhidas para a pesquisa. Mas primeiro temos que dizer uma coisa: em um espectro cromático, os limites entre uma cor e outra são

praticamente imperceptíveis, já que as cores adjacentes aos poucos vão se tornando a cor em questão. Mas o aspecto mais interessante é que essas 11 cores parecem dispor-se numa escala hierárquica implicacional:

Preto e/ou branco < Vermelho < Amarelo e/ou verde < Azul < Marrom < Púrpura e/ou rosa e/ou laranja e/ou cinza

Há línguas que possuem somente duas palavras para indicar cores. Nesse caso as cores só podem ser preto e branco (jalé, língua da Nova Guiné); quando aparece um terceiro termo, esse tem que corresponder necessariamente ao vermelho (tiv, língua niger-congo da Nigéria); se aparece um quarto e um quinto termo, esses podem ser amarelo e/ou verde (hanunóo, língua austroasiática das Filipinas); com um sexto termo se tem o azul (tamil, língua dravídica da Índia); o sétimo deve ser o marrom (como em nez perce, língua ameríndia do grupo penuti nas regiões norte da América) e assim vai.

A estrutura implicacional impõe que não pode se ter acesso a um nível sem ter os anteriores. Portanto é impossível que uma língua tenha o azul sem ter vermelho, amarelo e verde. Há línguas, como o português, que dispõem de mais de 11 termos para diferenciar as cores. Nesse caso isso implica que 11 são necessariamente os que fazem parte da hierarquia e não foram observadas tendências relevantes interlinguisticamente para os que vão além dos primeiros 11. Ou seja, a partir da posição 12 não foi notada alguma ordem hierárquica. Os tipos possíveis com base nas 11 cores são cerca de 2000, mas somente 22 são atestados. Isso deve significar algo cognitivamente.