

Flávia Santos Duarte de Oliveira

Eduardo Kenedy

Conselho Editorial

Arálba Teixeira de Castilho
Carlos Eduardo Lins da Silva
José Luiz Florin
Magda Soares
Pedro Paulo Funari
Rosângela Doin de Almeida
Tania Regina de Luca

Proibida a reprodução total ou parcial em qualquer mídia
sem a autorização escrita da Editora.
Os infratores estão sujeitos às penas da lei.

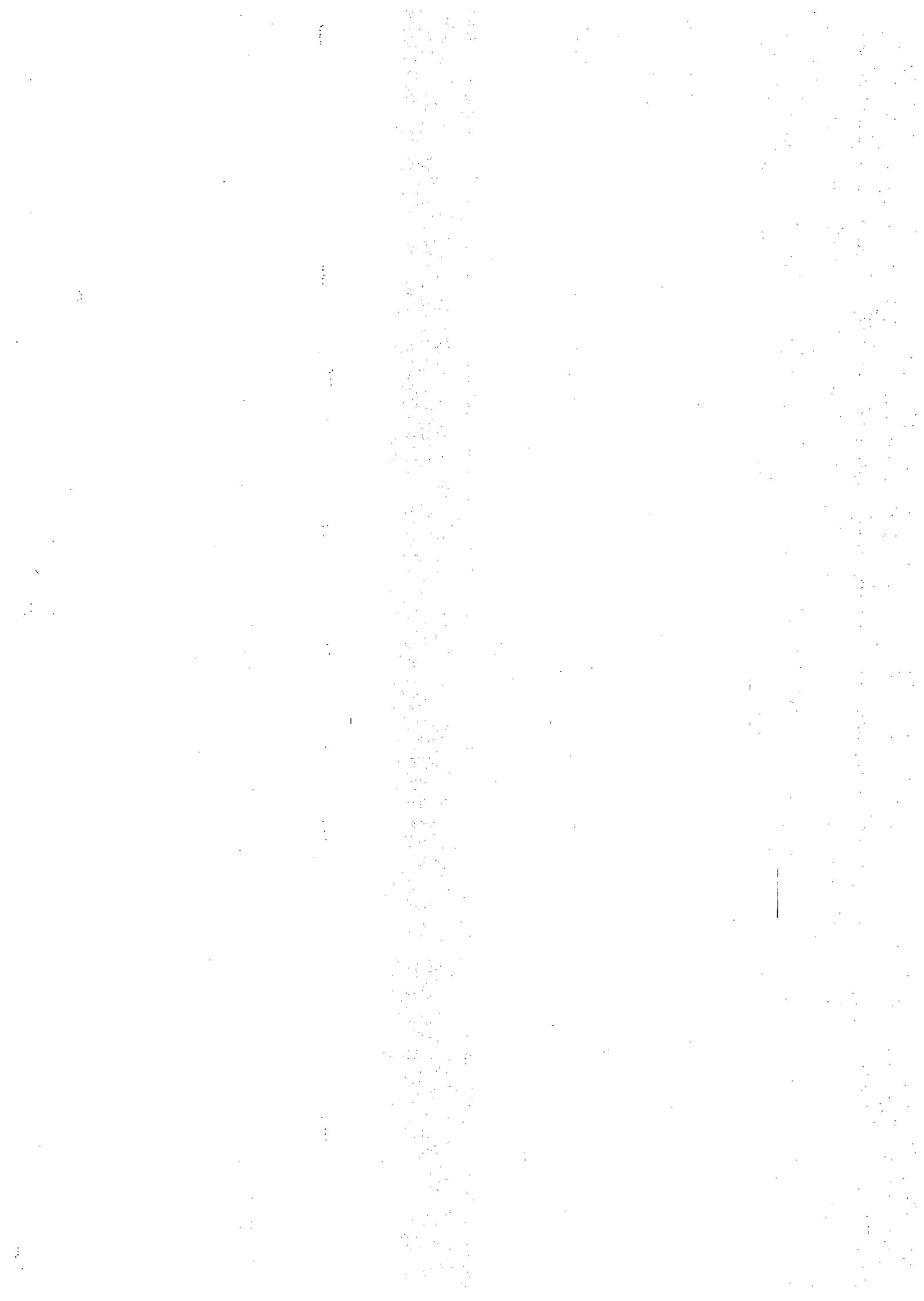
A Editora não é responsável pelo conteúdo da Obra,
com o qual não necessariamente concorda. O Autor conhece os fatos narrados,
pelos quais é responsável, assim como se responsabiliza pelos juízos emitidos.

Consulte nosso catálogo completo e últimos lançamentos em www.editoracontexto.com.br

Curso básico de linguística gerativa



editoracontexto



As referências bibliográficas, no corpo do texto das unidades, foram reduzidas ao mínimo indispensável. Essa opção motivou-se pelo desejo de conferir ao curso a sensação de um diálogo sincero, num simulacro do espaço real da sala de aula. Ao final do livro, porém, apresentam-se todas as obras que direta ou indiretamente foram citadas na apresentação dos conteúdos.

A articulação entre os princípios teóricos da linguística gerativa e as pesquisas empíricas da psicolinguística e da neurociência da linguagem faz-se presente em todos os momentos do curso. Essa triangulação visa explicitar ao estudante o caráter interdisciplinar do empreendimento gerativista contemporâneo, que se caracteriza como um programa de pesquisa maduro no contexto das ciências cognitivas.

O autor expressa profundo agradecimento a todos aqueles que de alguma forma contribuíram na elaboração deste curso. Em especial, agradece aos seus alunos e orientandos que, há mais de uma década, vêm formulando dúvidas e críticas ao Programa Minimalista e que neste livro encontram material didático em resposta a seus questionamentos. Também é grato aos professores Márcio Leirão (URPB), Ricardo Lima (UERJ), Marcus Maia (URJ), Mercedes Marciliese (URFP), Humberto Menezes (URJ), José Ferrari-Neto (URPB), Gabriel Othero (URRGS), Antônio Ribeiro (Faetec), Gerson Rodrigues (URRJ), Elisângela Teixeira (UFC) e Karine Vieira (CEDERJ) pelas preciosas sugestões feitas ao original da obra. A equipe de revisores e designers do CEDERJ e aos orientandos Maurício de Carvalho, Juliana Benevides e Luiz Guimarães, agradece pela cuidadosa revisão textual e pela participação na produção de inúmeras figuras presentes neste livro.

UNIDADE 1

A cognição linguística

Como estudante universitário, você talvez já tenha direcionado sua curiosidade intelectual para as seguintes questões linguísticas:

- Todos os seres humanos, exceto aqueles acometidos por alguma grave patologia, possuem a faculdade de produzir e compreender expressões linguísticas nas inúmeras situações do cotidiano que envolvem comunicação através de uma língua natural, como o português, o espanhol, o inglês ou qualquer outra. *O que é, afinal, essa faculdade humana para a linguagem?*
- Bebês não demonstram, ao nascimento, capacidade de produzir e compreender palavras, frases ou discursos, mas, ao longo de um tempo muito curto, não superior a três anos, essa faculdade já se manifesta de maneira bastante produtiva. Por volta dos cinco anos, uma criança já demonstra habilidade linguística equivalente à de adultos. *O que acontece com a criança durante o período em que ela está adquirindo a língua de seu ambiente?*
- Já na adolescência, a capacidade de adquirir uma língua de maneira natural decai significativamente. A partir de então, aprender uma nova língua demanda esforços conscientes que não são necessários durante a aquisição da linguagem em tenra infância, tais como frequentar cursos, ler manuais didáticos e dicionários, treinar a fala, corrigir erros com ajuda de professores etc. *Por que aprender línguas estrangeiras é tão diferente de adquirir uma língua-mãe?*
- Pessoas que sofreram derrames cerebrais ou que são acometidas por doenças neurológicas graves podem perder a capacidade linguística parcial ou totalmente. Na verdade, mesmo indivíduos sem queixas neurológicas podem ter o uso da linguagem comprometido quando se encontram desconcentradas ou muito cansadas. *Como é que o cérebro humano faz emergir as nossas capacidades linguísticas normais?*

Se você se interessou pelos problemas científicos que essas questões apresentam, isso significa que *linguagem e cognição* lhe parecem temas de pesquisa relevantes. Seja, portanto, bem-vindo ao nosso curso! Você tem em suas mãos um livro que pretende descrever e explicar os fundamentos da linguística gerativa — uma ciência da linguagem dedicada à dimensão cognitiva das línguas humanas.

Cognição é o termo científico atualmente utilizado para fazer referência ao conjunto das inteligências humanas. Diz respeito, portanto, a todos os fenômenos mentais que tenham relação com a aquisição, o armazenamento, a ativação e o uso de *conhecimento*. Conhecimento, por sua vez, é a palavra que usamos para fazer referência aos estados cognitivos de uma pessoa, os quais resultam da interação do indivíduo com o seu ambiente físico e sociocultural.

Ao longo das dez unidades deste curso, caracterizaremos a linguagem humana como sistema cognitivo. Colocaremos as línguas sob o microscópio da linguística para melhor compreendermos a natureza e o funcionamento do conhecimento linguístico humano. Começaremos pela presente unidade, em que apresentaremos a você a linguagem enquanto fenômeno cognitivo e a linguística gerativa como uma das ciências da cognição. Você é nosso convidado especial nesta introdução em dez etapas. Faremos uma breve jornada pela mente humana, num de seus nichos mais espetaculares: a linguagem e as línguas naturais. Prepare-se, pois a nossa viagem já começou.

A linguagem como fenômeno cognitivo

A linguagem humana, instanciada numa *língua natural*, é um fenômeno impressionante. Por meio de algumas dúzias de sons, podemos produzir e compreender palavras, frases e discursos que expressam os nossos pensamentos e que permitem o entendimento dos pensamentos das outras pessoas. Na verdade, tais sons podem ser substituídos por sinais entre os surdos ou por letras na língua escrita sem que o poder mobilizador da linguagem seja significativamente alterado.

Uma língua natural — como o português, o xavante, o alemão ou qualquer outra dentre as mais de seis mil línguas existentes hoje no mundo — é aquela que emergiu de maneira espontânea e não deliberada no curso da história humana. Opostas às línguas naturais, figuram as línguas artificiais. Uma língua artificial — como o esperanto e aquelas criadas em obras de ficção ou em códigos de computador — são conscientemente inventadas por uma pessoa ou por um grupo de indivíduos.

Imagine a seguinte situação. Uma pessoa apressa-se pelas ruas movimentadas do centro de uma cidade. Ela receia estar atrasada para um compromisso importante. Procurando informar-se sobre as horas, dirige-se a um transeunte e lhe diz: “Com licença, o senhor pode me informar as horas?”. O transeunte, por sua vez, compreende o estado mental de seu interlocutor e busca o comportamento adequado para a situação: olha o relógio de pulso, retira dele a informação necessária e produz a frase-resposta “São dez e meia”. Um evento trivial como esse esconde sob si um acontecimento extraordinário: o funcionamento da mente humana na tarefa de produzir e compreender expressões linguísticas numa língua natural.

Embora raramente tenhamos consciência disso, *a pessoa que produz frases e discursos* envolve-se numa tarefa psicológica extremamente complexa. Ela precisa codificar seus pensamentos em morfemas e palavras, que, por sua vez, devem ser combinadas entre si em sintagmas e frases, as quais, por fim, têm de ser pronunciadas para um interlocutor num dado contexto discursivo. Da mesma forma, *a pessoa que compreende frases e discursos* executa um trabalho mental igualmente engenhoso. Ela precisa decodificar os sons da fala que lhe são dirigidos no ato do discurso, de modo a identificar morfemas, palavras, sintagmas e frases para, assim, conseguir interpretar os estados mentais e as intenções comunicativas de seu colutor. Ora, podemos nos perguntar: como os humanos fazem isso? De que maneira essa sequência de codificação e decodificação de formas linguísticas ocorre? Onde ela acontece?

Apesar de ainda não sabermos tudo sobre como esses processos ocorrem, não há dúvida de que eles tenham lugar na mente humana. Dentre as diversas habilidades cognitivas de nossa espécie, a capacidade de produzir e compreender frases e discursos é uma das mais notáveis — e esse é o fenômeno mental oculto em nossas conversas cotidianas.

Figura 1.1: É na mente humana que a linguagem é produzida e compreendida.

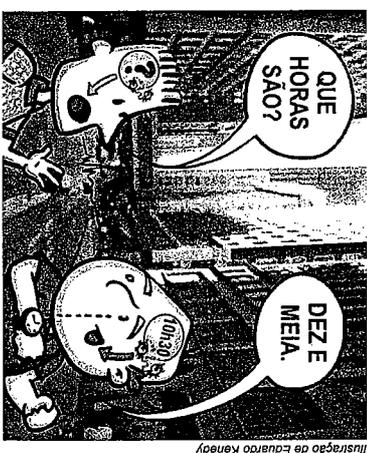


Ilustração de Eduardo Kenedy

As expressões linguísticas são as estruturas que ordenam o trânsito dos significados que vão de uma mente à outra entre indivíduos durante o discurso. São organizações silenciosas, das quais quase nunca tomamos consciência quando falamos ou ouvimos algo. De fato, os dois indivíduos da situação que citamos anteriormente estão preocupados com o conteúdo do que é dito. Eles sequer se dão conta de que é compulsório que esse conteúdo seja mentalmente codificado em estruturas complexas, como sintagmas e frases. O mesmo acontece com todos os humanos em todas as situações comuns de uso da linguagem: muito raramente tomamos consciência das formas que usamos para expressar certos conteúdos. A linguagem é, portanto, um *conhecimento tácito, implícito, inconsciente* no conjunto da cognição humana. Denominamos esse tipo de conhecimento como *conhecimento linguístico* ou *competência linguística*.

Além de silenciosas, as estruturas das línguas naturais manipuladas pela mente são geralmente muito complexas. Mesmo se analisássemos uma frase simples como "O senhor pode me informar as horas?", encontraríamos nela regras de ordenação de palavras, concordância, regência, seleção de pronomes... Enfim, verificaríamos a existência de uma complicada maquinaria gramatical a serviço da veiculação de significados. Entretanto, a despeito de toda essa complexidade, nós humanos somos capazes de produzir e compreender frases e discursos com extrema facilidade. Numa conversa qualquer, produzimos e compreendemos dezenas, centenas, milhares de frases, uma atrás da outra, numa velocidade incrivelmente rápida, muitas vezes medida em milésimos de segundo. Em circunstâncias normais, fazemos isso de maneira inconsciente e sem esforço cognitivo aparente.

Tipos de conhecimento

No estudo da cognição, distinguiamos "conhecimento declarativo", do qual somos conscientes, e "conhecimento tácito", do qual não temos consciência. Por exemplo, se você é capaz de clar os afluentes do Rio Amazonas, esse é um tipo de conhecimento declarativo que possui. Já se você sabe andar de bicicleta, dificilmente conseguirá explicar quais são os detalhes dos movimentos que o seu corpo precisa executar para manter o equilíbrio sobre duas rodas. Esse é, então, um tipo de conhecimento tácito em sua mente. Com relação particularmente à linguagem, as informações que adquirimos na escola sobre a gramática da língua portuguesa – como, por exemplo, o nome das classes de palavras e das funções sintáticas – são uma espécie de conhecimento declarativo/explicito. Por outro lado, a nossa capacidade de produzir e compreender palavras, frases e discursos de maneira natural em nosso cotidiano é exemplo de conhecimento tácito/implícito. Note que esse tipo de conhecimento tácito já existia em você mesmo antes que começasse a frequentar escolas e ter aulas de língua portuguesa, bem como existe naqueles indivíduos que sabem produzir e compreender expressões linguísticas normalmente mesmo sem nunca terem frequentado escolas (os analfabetos ou semianalfabetos).

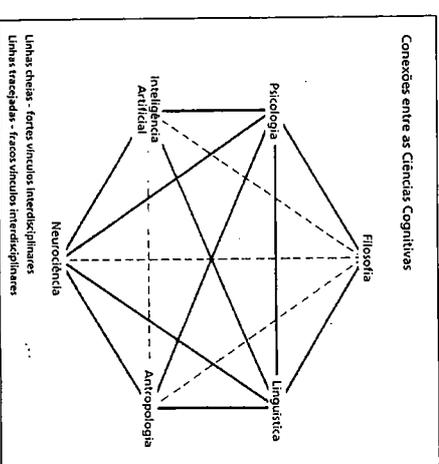
Como somos capazes disso? De que maneira nossas mentes tornam-se aptas a estruturar nossos pensamentos em sequências de sons ou sinais? Os resultados das pesquisas modernas nas ciências da cognição indicam que usamos a linguagem tal como o fazemos simplesmente porque somos seres humanos. Temos essa capacidade porque ela é uma característica natural de nossa espécie. *A facilidade da linguagem é, com efeito, a disposição biológica* que todos os indivíduos humanos saudáveis possuem para adquirir uma língua e para produzir e compreender palavras, frases e discursos.

Não obstante, tão importante quanto entendermos *por que* possuímos a facilidade da linguagem é descobrir *como* é a sua natureza e o seu funcionamento na mente dos indivíduos. *O porquê* e o *como* da linguagem na mente humana são os objetos de pesquisa da linguística enquanto uma das ciências cognitivas.

A linguística como ciência cognitiva

Chamamos de *ciências cognitivas* o conjunto das disciplinas que têm em comum o objetivo de compreender a natureza e o funcionamento da mente humana, a nossa cognição. Há um grande número dessas ciências, como a psicologia, a neurociência, a inteligência artificial, a filosofia da mente, a antropologia. Tais ciências são independentes, mas travam entre si profundo e intenso diálogo interdisciplinar. Dentre seus objetos de estudo figuram diferentes fenômenos cognitivos como visão, atenção, memória, aprendizagem, conceitos e categorias, raciocínio e dedução, resolução de problemas, tomada de decisão e muitos outros.

Figura 1.2: Algumas das ciências cognitivas.



Um dos fenômenos mentais mais importantes no estudo da cognição humana é a linguagem, e é justamente a ela que se dedica a linguística. Isso quer dizer que, quando os linguistas interpretam a linguagem como uma faculdade psicológica dos seres humanos, a linguística passa a integrar as ciências cognitivas.

Na qualidade de ciência cognitiva, a especificidade da linguística é evidente: ela tem a tarefa de descrever e explicar a *natureza, a origem e o uso* da linguagem humana. Diante desses objetivos, o trabalho do linguista pode ser sumarizado como um esforço para encontrar respostas para as seguintes perguntas:

1. Em que consiste o conhecimento linguístico existente na mente das pessoas?
2. Como esse conhecimento é adquirido pela criança já nos primeiros anos de vida?
3. Como esse conhecimento é posto em uso, em situação real, pelos indivíduos?
4. Como esse conhecimento é produzido pelo cérebro humano?

Na busca de respostas para tais questões, a linguística subdivide-se em três áreas especializadas em certos tipos de problemas: a *teoria linguística*, a *psicolinguística* e a *neurolinguística*, disciplinas que abordaremos a seguir.

Teoria linguística

A teoria linguística ocupa-se da questão 1: “em que consiste o conhecimento linguístico existente na mente das pessoas?” Sua tarefa é, portanto, formular uma teoria que explicite a natureza do conhecimento linguístico inscrito na mente dos seres humanos. Ou seja, o linguista teórico deve elaborar uma hipótese abstrata a respeito de como a linguagem deve funcionar na mente humana. Por exemplo, ele pode propor que a linguagem possua diversos componentes – tais como o léxico, a sintaxe, a fonologia – e dizer que cada um deles seja altamente especializado num tipo específico de conhecimento – como o conhecimento sobre as palavras (o léxico), o conhecimento sobre a estrutura das frases (a sintaxe) e o conhecimento sobre os sons (fonologia) da língua. O linguista teórico apenas cria hipóteses como essas. Seu trabalho é, essencialmente, um trabalho de abstração: formular hipóteses sobre como deve ser o conhecimento linguístico existente na mente das pessoas.

Nas ciências da cognição, existem diversos modelos que procuram articular respostas para a nossa questão 1. Esses modelos formulam, deste modo, teorias sobre como é o conhecimento linguístico humano. O modelo linguístico mais

influente nas ciências cognitivas é o *gerativismo*, que também pode ser chamado de *linguística gerativa*, *gramática gerativa*, *teoria gerativa*, ou ainda, como era chamado num termo mais antigo, *gramática gerativo-transformatacional*. Nosso curso apresentará a você os fundamentos do gerativismo, que será aqui convencionalmente interpretado como termo mais ou menos sinônimo de “teoria linguística” – mas devemos estar conscientes de que existem outros modelos na teoria linguística que não são gerativistas. Ao final das dez unidades de nosso curso, caso tenha interesse em prosseguir com os estudos sobre o conhecimento linguístico humano, você estará apto a entender como os outros modelos diferenciam-se do gerativismo e formulam teorias próprias e independentes.

No uso informal e cotidiano da linguagem, o termo “teoria” geralmente significa “opinião”, “hipótese não comprovada” ou mesmo “palpite”. Em ciência, “teoria” assume outro significado, bastante específico e muito diferente de seu valor popular. Nas ciências, uma “teoria” é uma explicação unificada para um conjunto de dados e de observações sobre determinado fenômeno. Por exemplo, na biologia, a seleção natural é a “teoria” que explica unificadamente os fatos da evolução orgânica; na física, a “teoria” gravitacional explica de maneira unificada os fenômenos da queda dos corpos na Terra e do movimento dos planetas no universo. Uma “teoria linguística” deve, portanto, prover explicações sobre os fatos da linguagem na mente e no cérebro dos seres humanos.

Em ciência, um modelo é um conjunto de hipóteses sobre a estrutura e o funcionamento de algum fenômeno – no nosso caso, a linguagem humana. A função de um modelo é formular uma teoria que explique e preveja as propriedades do fenômeno a que se dedica.

O gerativismo teve início nos anos 1950 do século XX, quando Noam Chomsky, norte-americano, professor do MIT (Instituto de Tecnologia de Massachusetts, EUA), formulou suas primeiras ideias a respeito da natureza mental da linguagem humana. Já em seu primeiro livro (*Estruturas Sintáticas*, publicado em 1957), Chomsky afirmava que o papel fundamental da linguística é tornar explícito, isto é, descrever com objetividade científica, o conhecimento linguístico dos falantes. Para ele, a teoria linguística deve descrever os procedimentos mentais que “geram” as estruturas da linguagem, como as palavras, as frases e os discursos. Por exemplo, ele mesmo levantou a hipótese (isto é, criou uma teoria) segundo a qual as frases são criadas na mente das pessoas por meio de aplicações de regras inconscientes, as quais se aplicam sobre certas palavras de modo a “gerar” as frases que pronunciamos e compreendemos.

Noam Chomsky

Chomsky nasceu nos EUA, em 1928. Além de linguista, é filósofo e ativista político. Ao longo de mais de 60 anos de sua vida acadêmica, publicou dezenas de livros e centenas de artigos científicos. Chomsky é considerado um dos pensadores mais importantes da história moderna. Suas ideias revolucionaram o estudo da linguagem e inseriram a linguística no contexto da revolução cognitiva dos anos 60 do século XX. Chomsky foi um severo crítico da psicologia behaviorista dominante na primeira metade do século passado. Para os behavioristas mais radicais, todos os tipos de comportamento humano ou animal são gerados externamente, por meio de cadeias associativas entre dados estímulos e certas repostas. Para esses estudiosos, a associação entre estímulo e resposta é criada pela repetição, por meio de “recompensas” ou “reforços” advindos do ambiente. Segundo um behaviorista, o aprendizado pela pura repetição aconteceria mesmo no que diz respeito à linguagem humana, por ele denominada “comportamento linguístico”. Em 1959, Chomsky publicou sua clássica resenha sobre o livro *Comportamento Verbal*, do famoso behaviorista B. F. Skinner. Na resenha, Chomsky demonstrou o caráter criativo da linguagem humana, sua natureza mental e abstrata, por oposição ao modelo de linguagem como “comportamento condicionado pelo ambiente” defendido pelos behavioristas. Desde então, Chomsky vem empenhando-se na formulação de uma teoria sobre a natureza da linguagem na mente humana. No estágio atual de sua pesquisa, o chamado Programa Minimalista, Chomsky defende a hipótese de que todas as línguas naturais são um conjunto de Princípios universais e inatos e de Parâmetros, também inatos, que são formatados durante o período da aquisição da linguagem.

A abordagem de Chomsky foi revolucionária para a época, pois, até a metade do século passado, a linguística ocupava-se quase exclusivamente da dimensão social e histórica da linguagem humana, tal como acontecia no estruturalismo linguístico. A partir das ideias de Chomsky, os linguistas passaram a não apenas descrever a estrutura das línguas, mas também a procurar explicações para como a mente humana era capaz de adquirir e processar essas estruturas. Com Chomsky, a morada da linguagem e das línguas naturais passou a ser a mente dos indivíduos.

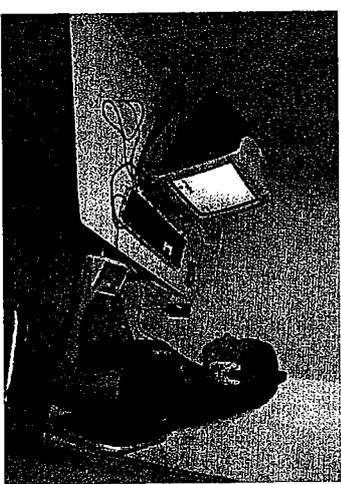
Veremos, ao longo das unidades de nosso curso, como o gerativismo vem formulando teorias que procuram responder à pergunta “o que é conhecimento linguístico?”. Ou seja, veremos como a teoria linguística vem respondendo à questão 1.

Psicolinguística

Enquanto o gerativismo cumpre a função de formular teorias abstratas sobre a linguagem como fenômeno cognitivo, a psicolinguística caracteriza-se como uma

ciência empírica cujo objetivo é investigar de que maneira as crianças adquirem uma língua natural e como os indivíduos adultos produzem e compreendem palavras, frases e discursos no tempo real da comunicação cotidiana. A psicolinguística procura encontrar, portanto, respostas para as questões 2 e 3: “como as crianças adquirem o conhecimento linguístico?”, “como esse conhecimento é colocado em uso?”.

Figura 1.3: Iovem participa de experimento psicolinguístico.



A psicolinguística é uma ciência empírica, que utiliza os métodos e as técnicas experimentais da **psicologia cognitiva** como instrumento de pesquisa acerca da aquisição e do uso das línguas naturais. Como é possível imaginar, a psicolinguística pode ser uma disciplina fortemente inter-relacionada à teoria linguística no conjunto das ciências cognitivas. Dentre suas diversas ocupações científicas, a psicolinguística pode aliar-se ao gerativismo para testar empiricamente as previsões dos modelos abstratos formulados pelos gerativistas. Por sua vez, a teoria linguística pode basear-se nas descobertas da psicolinguística para formular ou reformular suas hipóteses teóricas, construindo, dessa maneira, modelos que se sustentem na realidade psicológica do funcionamento real da linguagem.

A psicologia cognitiva é uma das áreas mais recentes das pesquisas em psicologia, tendo surgido ao final dos anos 1950, com a revolução cognitiva. Trata-se da área da psicologia que procura estudar, através de experimentos científicos, a natureza e o funcionamento da cognição humana. Um psicólogo cognitivo dedica-se a fenômenos como, dentre outros, memória, atenção, percepção, raciocínio, resolução de problemas, linguagem, emoções. A psicolinguística é uma das subdisciplinas da psicologia cognitiva.

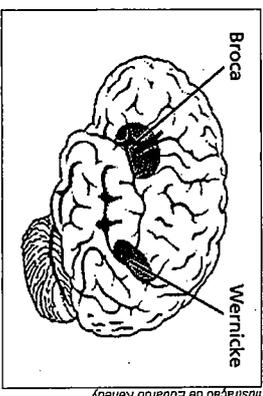
Neurolinguística

A **neurolinguística** é uma das subdivisões das neurociências modernas. Trata-se de uma ciência empírica cujo objetivo é compreender os mecanismos cerebrais que dão origem à linguagem humana. Enquanto a psicolinguística dedica-se ao estudo da mente, isto é, das funções cognitivas visíveis no comportamento humano, a neurolinguística ocupa-se do cérebro, seus neurônios e suas sinapses – os sistemas físicos, químicos e biológicos que dão origem à mente.

Atualmente, prefere-se a expressão “neurociência da linguagem” em vez do tradicional termo “neurolinguística”. O novo termo é útil, inclusive, para evitar confusão com a “programação neurolinguística”, que não possui relação com a neurolinguística que apresentamos nesta unidade.

Uma ilustração útil para compreendermos a diferença entre “mente *versus* cérebro” e “psicolinguística *versus* neurolinguística” é a analogia com os computadores digitais modernos. Praticamente todos nós sabemos que os computadores possuem duas partes fundamentais: (1) os *softwares*, que são os programas que usamos para escrever textos, fazer cálculos, ouvir músicas e navegar na internet, e (2) o *hardware*, a parte física do computador que torna o seu uso possível, como o teclado, o mouse, o monitor, o disco rígido etc. Se fizermos uma metáfora, entenderemos que a mente são os nossos *softwares* psicológicos, como a linguagem, a visão, o raciocínio etc., enquanto a superfície física que torna possível o uso desses *softwares* é o cérebro, o nosso *hardware* neuronal.

Figura 1.4: Áreas do cérebro associadas à linguagem humana: área de Broca (produção) e área de Wernicke (compreensão).



Pelo que dissemos, podemos entender que a neurolinguística é a ciência que deve encontrar respostas para a quarta questão que formulamos: “quais são os fundamentos físicos, no cérebro humano, do conhecimento linguístico?”.

Três realidades no estudo da linguagem

Com base no que acabamos de dizer sobre a distinção entre teoria linguística, psicolinguística e neurolinguística, é possível compreendermos que cada uma dessas áreas possui o seu próprio objeto de estudo no complexo fenômeno da linguagem na mente humana. Cada qual desempenha papel singular no desenvolvimento da linguística como ciência cognitiva. Os objetos particulares com que se ocupa cada uma das três disciplinas constituem a *realidade específica* do seu domínio de investigação.

Assim, a *realidade teórica* ou *epistemológica* do conhecimento linguístico é o objeto das pesquisas em teoria linguística. Já a psicolinguística tem como objeto de estudo a *realidade psicológica* das línguas naturais no seu funcionamento em tempo real na mente humana. Por fim, o objeto da neurolinguística é a *realidade neurológica* da linguagem no cérebro, em sua substância eletroquímica.

A realidade teórica do conhecimento linguístico constitui uma *abordagem cognitiva moderada* acerca da linguagem na mente humana, por oposição à *abordagem forte* da realidade psicológica e à *abordagem fortíssima* da realidade neurológica.

Dizemos que a teoria linguística possui uma abordagem cognitiva moderada porque deve ocupar-se primordialmente com o *quê* da competência linguística humana, e não com o *como* dos processos mentais e cerebrais a ela subjacentes. O teórico da linguagem não necessariamente frequenta laboratórios científicos ou realiza trabalhos empíricos que observam o comportamento da mente e do cérebro em funcionamento real. Tal como Chomsky, o linguista teórico pode trabalhar exclusivamente em seu escritório, formulando teorias de maneira conceitual e dialética, bem ao estilo da filosofia.

Por sua vez, a psicolinguística deve ser caracterizada como uma abordagem cognitiva forte. Isso acontece porque os psicolinguistas observam e descrevem empiricamente, com base em pesquisas experimentais realizadas em laboratórios, os processos psicológicos pelos quais o conhecimento linguístico se realiza na mente humana. Os psicolinguistas idealizam e executam experimentos científicos que envolvem tarefas cognitivas observáveis como, por exemplo, a produção e a compreensão de palavras, frases e discursos. A partir dos resultados de experimentos desse tipo, os psicolinguistas podem formular generalizações a respeito de como a mente humana processa a linguagem natural.

A realidade neurológica da pesquisa neurolinguística caracteriza-se, por fim, como uma abordagem cognitiva fortíssima em razão de, em suas pesquisas, o neurolinguista considerar não somente o *quê* abstrato e o *como* mental do conhe-

cimento linguístico, mas também o *onde* dos processos neurológicos que realizam fisicamente a linguagem na substância neuronal do cérebro.

A distinção entre as três realidades no estudo cognitivo da linguagem é importante porque a realidade epistemológica da teoria linguística é relativamente independente da realidade psicológica e neurológica. Por exemplo, quando um linguista teórico apresenta um conceito como “sintagma”, ele está formulando uma teoria que explica como a mente humana relaciona as palavras umas às outras no interior de uma frase. Ele não está dizendo que encontraremos “sintagmas” se abirmos o crânio de uma pessoa e analisarmos a configuração dos neurônios na superfície de seu cérebro.

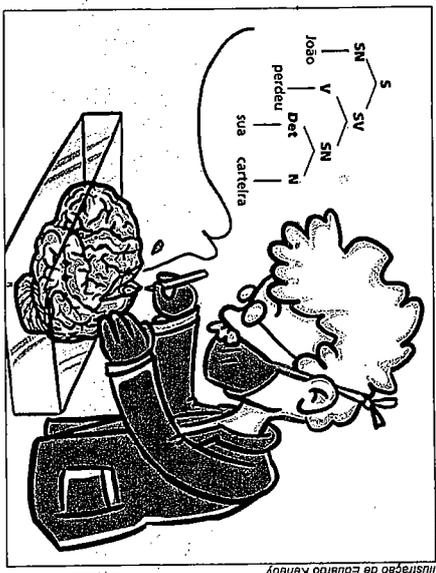


Figura 1.5: As relações entre as três realidades da linguagem não são imediatas ou transparentes.

A integração entre as três realidades

Para o progresso do estudo da linguagem como fenômeno cognitivo, teoria linguística, psicolinguística e neurolinguística devem manter entre si intensas e profundas relações no empreendimento das ciências cognitivas. Por um lado, a teoria linguística deve formular modelos abstratos que representem o conhecimento linguístico existente na mente das pessoas. Por outro lado, psicolinguística e neurolinguística podem dedicar-se a verificar de que maneira tais modelos articulam-se ou não com a realidade concreta do que se passa na mente e no cérebro dos indivíduos. Além disso, as descobertas da psicolinguística e da neurolinguística podem orientar os linguistas teóricos em sua busca por um modelo cientificamente preciso, que represente de maneira correta o conhecimento linguístico humano.

No melhor dos cenários do desenvolvimento da linguística como ciência cognitiva, os estudos da teoria linguística serão articulados às descobertas sobre a realidade psicológica e neurológica da linguagem, de modo que se torne possível formular respostas integradas às questões “o que é conhecimento linguístico?”, “como ele é adquirido?”, “como ele é usado?” e “quais são seus substratos neurológicos?”. No futuro do estudo da linguagem na mente, deverá haver uma integração cada vez maior entre as realidades epistemológica, psicológica e física do conhecimento linguístico.

Conclusão

Nesta unidade, aprendemos que a linguagem é um tipo de conhecimento e, assim, faz parte da cognição humana. Como tal, a linguagem e as línguas naturais são objeto de estudo das ciências cognitivas, especialmente da teoria linguística, da psicolinguística e da neurolinguística. Vimos que, na teoria linguística, o gerativismo é o modelo mais influente nos últimos cinquenta anos. Noam Chomsky é o principal teórico e o precursor desse tipo particular de linguística. Aprendemos que, no futuro das ciências cognitivas, teoria linguística, psicolinguística e neurolinguística devem articular respostas integradas para as quatro questões fundamentais do estudo da linguagem como faculdade cognitiva humana.

Exercícios

- 1) Por que, dentre suas diversas características, a linguagem humana deve ser considerada um fenômeno cognitivo?
- 2) Visite o site “YouTube” e assista ao vídeo situado no seguinte endereço: http://www.youtube.com/watch?v=k-T_6WARZlQ. Após assistir ao vídeo, responda: quais são as relações entre filosofia, psicologia cognitiva e neurociência no estudo da linguagem humana?
- 3) Nas ciências cognitivas, qual é a função de uma teoria linguística?
- 4) O que é o gerativismo? Quem é Noam Chomsky?
- 5) Caracterize teoria linguística, psicolinguística e neurolinguística. Explique por que a realidade linguística a que cada uma dessas disciplinas se dedica é relativamente independente das demais e como é importante, para o progresso da ciência, que tais realidades sejam integradas no futuro.

UNIDADE 2

Conceitos fundamentais

Imagine que você esteja num encontro internacional de estudantes oriundos de vários países. Como cada estudante possui sua própria língua nativa e há muitas línguas no mundo, a primeira abordagem entre eles é geralmente feita em inglês, que funciona hoje como uma espécie de língua de comunicação universal. Em inglês, os estudantes apresentam-se entre si e conversam sobre o seu país de origem, sobre as línguas que sabem falar etc.

Se, nesse encontro, você fosse abordado por alguém que lhe perguntasse *qual* é a sua língua nativa, o que você responderia? Não há dúvida de que sua resposta seria rápida e fácil: você é um falante nativo do português.

E se lhe perguntassem *por que* você adquiriu especificamente o português e não outra língua como, por exemplo, o russo? Também aqui, sua resposta seria óbvia e imediata: você nasceu no Brasil, numa sociedade em que se fala o português, e não o russo, e foi nesse ambiente cultural lusófono que, em tenra infância, você adquiriu a linguagem.

Já se lhe perguntassem *onde* está a língua que você sabe falar, isto é, em que lugar se encontra o português, o que você diria? Nesse caso, sua resposta não seria assim tão fácil. Na verdade, a própria pergunta é um tanto incommum, de tal forma que o seu interlocutor poderia refazê-la com outras palavras: onde concretamente podemos localizar a língua portuguesa?

A estranheza da pergunta acontece por que, normalmente, não pensamos nas línguas como algo concreto, uma coisa que possa ser localizada da maneira como a pergunta pressupõe. No entanto, ainda que estranha, a questão é perfeitamente razoável, afinal, se existem no mundo real, as línguas devem ser localizadas no espaço e no tempo. Pense por um momento: onde no mundo concreto se encontram as línguas naturais?

Se você se lembra dos conteúdos da primeira unidade de nosso curso, não lhe restará dúvida de que uma língua se encontra na mente dos indivíduos que a falam. Essencialmente, é na mente das pessoas que uma língua natural encontra-se armazenada e é de lá que ela é ativada para o uso, nas tarefas cotidianas de produção e compreensão da linguagem. O português, por exemplo, encontra-se na sua mente (e na de todos os demais falantes dessa língua).

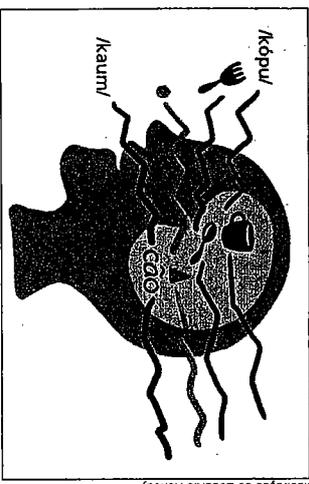


Figura 2.1: A mente humana é a morada da linguagem.

Mas, além dessa existência interior e subjetiva nos indivíduos, as línguas existem também no mundo objetivo exterior. Com efeito, para que uma língua natural se torne verdadeiramente útil e cumpra sua função primordial de organizar e comunicar pensamentos, é necessário que um grupo de indivíduos compartilhe, ainda que de forma aproximada e imperfeita, o código objetivo a ser manipulado pela linguagem. Dizendo de outra forma, para que uma língua natural exista de fato, é essencial que as pessoas compartilhem um léxico, e tudo o que nele está inscrito (fonemas, morfemas, palavras, expressões, significados e convenções de uso). É somente a partir de tais elementos compartilhados socialmente que a linguagem toma existência na mente de um indivíduo. Esse código objetivo compartilhado pelas pessoas é aquilo que compreendemos como a dimensão da linguagem externa à mente humana.

Considerando as duas dimensões da linguagem que citamos acima, isto é, a dimensão subjetiva e a objetiva, sua resposta à pergunta “onde se encontra a língua que você sabe falar” é, na verdade, dupla: por um lado, o português existe na sua mente e na de todos os indivíduos que falam essa língua e, por outro, existe também no código objetivo compartilhado pelas comunidades lusófonas, tais como as do Brasil, as de Portugal, as de Angola etc.

Se você compreendeu corretamente a dupla possibilidade de resposta à pergunta mencionada, então podemos aprofundar o problema na sequência desta unidade. Exploraremos aqui a dicotomia subjetividade e objetividade existente nas línguas naturais. Conforme analisaremos, não é apenas o português que possui uma dimensão linguística subjetiva e outra objetiva, mas todas as línguas humanas são igualmente bidimensionais. Veremos, nas próximas seções desta unidade, de que maneira essa dupla dimensão da linguagem é capturada pelos conceitos de língua-I e língua-E. Com base nesses conceitos, entendemos que é sobre a língua-I que recai o interesse de estudo do gerativismo. A compreensão do que é uma língua-I nos levará aos conceitos de modularidade da mente, modularidade da linguagem e interação entre módulos. Tais conceitos constituem os primeiros passos no longo caminho do estudo da linguagem na mente humana.

Conceitos de língua

Voltemos ao exemplo da introdução desta unidade. Você está num encontro internacional com estudantes de todas as partes do mundo. Os estudantes se comunicam inicialmente em inglês, que funciona como língua de comunicação universal. Ao ser perguntado sobre sua língua nativa, você responde que é falante do português. Então explica que essa foi a língua de seu ambiente na infância, no Brasil, quando você estava nos anos de aquisição da linguagem. Por fim, você declara que o português existe na sua mente, bem como na mente de todos os demais falantes dessa língua. Além disso, você afirma que o português existe também no código linguístico compartilhado por todas as comunidades de língua portuguesa ao redor do globo terrestre.

Nessa pequena história, quando formulamos a dupla resposta à pergunta “onde se encontra a língua que você sabe falar”, estávamos na verdade apresentando a você um problema crucial no estudo da linguagem: o conceito de língua. Como vimos, o termo *língua* pode assumir pelo menos dois significados. Primeiramente, pode significar o conhecimento linguístico de um indivíduo acerca de uma dada língua, ou seja, é a faculdade cognitiva que habilita esse indivíduo a produzir e compreender enunciados na língua de seu ambiente. Nessa acepção, o termo *língua* refere-se a uma habilidade presente na mente humana. Em segundo lugar, *língua* pode significar o código linguístico existente numa comunidade humana, isto é, língua é o léxico e tudo o que nele está contido ou dele é derivado. Nessa acepção, *língua* diz respeito a algo que assume existência fora da mente das pessoas. É com base no primeiro significado de *língua* que você pode responder à pergunta da história dizendo que “o português existe na sua mente e na mente de todos os que sabem falar essa língua”. E é a partir do segundo significado que faz sentido você responder à pergunta dizendo que “o português existe no código linguístico vigente nas comunidades lusófonas do mundo, tais como Brasil, Portugal, Moçambique etc.”

Você já deve ter percebido que *língua* como faculdade cognitiva e *língua* como código linguístico são realidades muito diferentes. Para exemplificar essa dessemelhança, pensemos, por um momento, na quantidade de línguas existentes no mundo. Quantas línguas você imagina que são faladas hoje ao redor do globo? Se assumirmos que *língua* é uma capacidade presente em cada mente humana, diremos que há no mundo 7 bilhões de línguas, na medida em que existem nele cerca de 7 bilhões de cérebros/mentes humanas. Já se interpretarmos *língua* como código linguístico socialmente compartilhado, então diremos que existem 6 mil línguas no mundo, cada qual com centenas de subdivisões dialetais correspondentes às diversas organizações sociais humanas. São números bastante diferentes, não são?

Dada a grande diferença entre os dois conceitos, você deve estar se perguntando como é possível evitar confusões ao empregar o termo *língua* ao longo de seus estudos sobre a linguagem humana. Afinal, como poderemos deixar claro quando estivermos falando de língua como faculdade cognitiva e quando estivermos falando de língua como código linguístico socialmente compartilhado?

Em seu livro clássico, lançado em 1986, intitulado *O conhecimento da língua: sua natureza, origem e uso*, Chomsky procurou resolver esse problema. Ele propôs o uso do termo “língua-I” para fazermos referência à língua em sua acepção cognitiva. Para nos referirmos à língua como fenômeno sociocultural, Chomsky propôs o termo “língua-E”.

Conceitos de Chomsky: língua-I e língua-E

[...] sobre] a existência da língua na mente do falante [...] Vamos referir-nos a esta noção como “língua interna (Língua-I). A língua-I é, pois, um elemento que existe na mente da pessoa que conhece a língua, adquirido por quem aprende e usado pelo falante-ouvinte” (1994: 41)

[...] uma língua é usada por uma população [...] por um interesse comunicativo. Vamos referir-nos a esse conceito como instância de uma “língua externa (Língua-E), no sentido em que o construir e compreenderdo independentemente das propriedades da mente/cérebro” (1994: 39)

Ao usarmos os conceitos propostos por Chomsky, poderemos distinguir, com clareza, a dimensão subjetiva das línguas e a sua dimensão objetiva, evitando, assim, as confusões que o termo genérico *língua* pode provocar. A dimensão mental/subjetiva do fenômeno da linguagem, que também chamamos de cognitiva ou psicológica, é sintetizada no conceito de “língua-I”, em que “I” significa interna, individual e **intensional** (escrito com “s” mesmo!). Por sua vez, a dimensão sociocultural/objetiva das línguas é denominada “língua-E”, em que “E” quer dizer externa e **extensional**.

A distinção intensional *versus* extensional tem origem na filosofia. “Intensional” diz respeito a tudo o que é inferior e próprio a um dado significado, ao passo que “extensional” refere-se à extensão de um dado significado, isto é, diz respeito à classe de objetos a que o significado se refere.

Se você já compreendeu corretamente que todo e qualquer fenômeno linguístico compreende sempre duas dimensões, a extensional (coletiva, exterior) e intensional (individual, interior), poderemos então passar às seções seguintes. Veremos em detalhes o que faz com que uma língua deva ser interpretada como língua-E e, por contraste, o que faz com que uma língua deva ser interpretada como língua-I. Você verá que, se especificarmos a noção de língua nos conceitos de língua-E e de língua-I, muitas confusões comuns nos estudos da linguagem poderão ser evitadas.

Língua-E

A noção de Língua-E corresponde *grossa modo* ao que comumente se interpreta como *língua* ou *ídioma* no senso comum, por pessoas que não frequentaram cursos especializados em linguística. Por exemplo, o *português* é uma língua-E no sentido de que é esse fenômeno *sociocultural, histórico e político* que compreende um *código linguístico*: a língua portuguesa.

Uma língua-E é um fenômeno *sociocultural* porque é compartilhada pelos indivíduos que integram uma mesma sociedade, com suas diversas nuances, e, dessa forma, compartilham uma cultura. Como sabemos, uma sociedade pode ser composta por um número muito pequeno ou extremamente grande de indivíduos, mas essa distinção numérica não é relevante para o conceito de língua-E. Pense, por exemplo, nos anambé, integrantes de uma sociedade indígena brasileira que vive na região central do Pará. Essa sociedade é composta por pouco mais de 100 indivíduos, os quais compartilham a língua anambé, que é a sua língua-E. Trata-se de uma pequena quantidade de falantes. Agora pense nas complexas nações multicontinentais modernas, que abrigam um gigantesco número de pessoas. A língua inglesa, por exemplo, é hoje falada por mais de 500 milhões de indivíduos que nasceram numa sociedade anglófona, como as existentes na Inglaterra, nos EUA, na Austrália, na África do Sul etc. O inglês é também usado por cerca de 1 bilhão de falantes não nativos, isto é, pessoas que aprenderam a língua na condição de falantes estrangeiros. Esse enorme contingente de falantes que se espalham por todos os continentes faz do inglês uma das línguas-E mais faladas no mundo. A despeito dessas diferenças numéricas e de prestígio internacional, anambé e inglês são o que chamamos de duas línguas-E.

Uma língua-E é um fenômeno *histórico* porque sempre se constitui ao longo do tempo, nas contingências da história da humanidade, em suas diversas populações ao redor do planeta. Por exemplo, a sua língua-E é o português, já que você é brasileiro, nascido numa comunidade que cultivou essa língua-E ao longo dos séculos de história do Brasil. Na verdade, o código linguístico

que você herdou ao nascer possui uma origem muito mais remota do que a da sua comunidade. Os fonemas, os morfemas, as palavras, os significados e usos do português no Brasil remontam não somente aos séculos passados da história da colonização brasileira, mas têm também origem nos séculos e milênios da história sociocultural de Portugal e do Império Romano, cuja língua, o latim, acabou dando origem ao português. Se quisermos ir ainda mais a fundo, poderíamos dizer que a história da sua língua-E se inicia antes mesmo da constituição do Império Romano, com as línguas pré-românicas que deixaram heranças ao latim, que as retransmitiu ao português. Note que você simplesmente herdou o português dessas circunstâncias históricas, sem sequer ter consciência delas. Se você tivesse nascido noutro lugar do mundo, ou noutro momento da história, sua língua-E seria outra. Shakespeare, por exemplo, herdou de seu ambiente uma língua germânica, o inglês da Inglaterra, do século XVI, que assim se tornou sua língua-E. Ao nascer, um ser humano herda um conjunto de contingências históricas e uma delas é a sua língua-E.

Por se tratar de um fenômeno sociocultural e histórico, uma língua-E é, por isso mesmo, um fato *político*. Com efeito, as sociedades desenvolvem-se historicamente por meio de ações políticas e muitas delas voltam-se para a linguagem. Por exemplo, a Constituição Federal Brasileira, de 1988, afirma em seu art. 13º que a língua oficial de nossa nação é o português. Isso é uma posição política perante a história do Brasil e, em particular, perante a história de nossa cultura. Numa outra postura política, um tanto radical, seria possível argumentar que nossa língua-E não é mais o português, mas, sim, um novo código, originado das mudanças que o português europeu sofreu na América, em contato com línguas indígenas e africanas: o *brasileiro*. Outro exemplo de atitude política na linguagem é promover uma língua-E institucionalmente, através do ensino formal em escolas e do cultivo de aprendizagem da língua como idioma estrangeiro.

Atualmente, vivemos no Brasil um problema político importante em relação à nossa língua-E: o tratamento dos estrangeirismos. Há um projeto de lei no Congresso Nacional que visa proibir o uso de termos estrangeiros, sobretudo os de origem inglesa, no comércio e na propaganda. Sem nos ocuparmos do mérito da questão, o importante para o contexto de nossa unidade é você entender que usar uma expressão como “fazer download” ou “baixar um arquivo” é uma postura política em relação à língua-E, que tem a ver com o nacionalismo e a negação de estrangeirismos. Qual é a sua opinião? Você é contra ou a favor dos estrangeirismos? Acha que devem ser mesmo proibidos por lei? Qual é a sua opinião política sobre sua língua-E?

Você deve estar se perguntando qual é a importância de uma língua-E para uma ciência de cunho cognitivista como a linguística gerativa. Bem, podemos dizer

que, para um gerativista, uma língua-E é simplesmente um código linguístico, um léxico com seus inúmeros componentes. Esse léxico compreende as palavras compartilhadas pelos membros da coletividade para fazer referências às mais distintas entidades de seu universo físico e cultural. Compreende, da mesma forma, especificações quanto aos sons usados na comunidade de fala (fones e fonemas), especificações quanto às categorias morfosintáticas consideradas relevantes para a língua, tais como, dentre outras, classes de palavras, gêneros gramaticais, tempos, modos, aspectos e pessoas verbais. E, finalmente, compreende especificações quanto às convenções de uso, como formas de tratamento, usos elegantes ou ofensivos, linguagem figurada, expressões frequentes, expressões raras, dentre muitos outros tipos de informação.

Léxico e cognição linguística

Nos estudos linguísticos, o léxico não é somente o vocabulário de uma língua, não é apenas o conjunto de palavras convencionalizadas numa língua-E, tal como um dicionário. O léxico é, sobretudo, a maior fonte de informação linguística necessária para a aquisição e o uso de uma língua-I. Assim, o léxico contém, é claro, as especificações sobre a relação arbitrária entre o significante e o significado de um grande número de palavras – tal propôs Ferdinand de Saussure, o famoso linguista considerado precursor da linguística moderna –, mas contém o conjunto de todas as informações fonológicas, morfosintáticas e semânticas imprescindíveis para a aquisição e o uso da língua do ambiente. Por exemplo, é no conjunto de informações codificadas no léxico do português (língua-E) que estão especificados os traços linguísticos que compõem o item “casa”, isto é, a sua pronúncia, com sílabas e fonemas específicos, sua classe gramatical (nome), seu gênero gramatical (feminino), seu significado (lugar destinado à moradia), dentre outras informações. Veremos mais sobre a noção cognitiva de léxico na unidade 8 de nosso curso.

Para os estudiosos da linguística gerativa, o interesse de estudo acerca das línguas-E concentra-se na descrição das informações que estão codificadas no léxico dessas línguas. Essas informações são bastante variáveis de língua para língua, já que, como sabemos, uma língua-E é um produto sócio-histórico, e, por isso mesmo, está sujeita às contingências da experiência cultural humana. O interesse particular do gerativista recai sobre o fato de a mente humana ser capaz de adquirir essas informações, sejam quais forem, para, a partir delas, produzir e compreender expressões linguísticas no uso cotidiano da língua-I.

Dessa forma, o gerativista procurará identificar, ao descrever uma língua-E, os traços linguísticos que estão codificados em seu léxico e são utilizados na formação de representações mentais, como fonemas, morfemas, palavras, sintagmas, frases e discursos.

Crítérios políticos na identificação de línguas

Para ilustrarmos de que maneira a noção de língua-E é influenciada por fatores políticos, pensemos nas línguas faladas na China e na Escandinávia. Imaginamos que o chinês seja uma e apenas uma língua, mas os falantes chineses sabem que a língua que se fala em Beijing (dialeto mandarim) é totalmente diferente daquela falada em Hong Kong (dialeto cantonês), a ponto de a comunicação oral entre os falantes dessas duas comunidades se tornar impossível, a não ser que se recorra à tradução ou à escrita ideogramática comum aos dois dialetos. Ora, mandarim e cantonês são claramente dois códigos linguísticos muito diferentes. Por que então não dizemos que a China possui duas línguas? Não dizemos isso porque a unidade geopolítica daquele vasto país é reforçada com a ideia de que uma única língua é usada por todos os seus bilhões de habitantes. Na Escandinávia, região europeia que abrange a Noruega, a Dinamarca e a Suécia, acontece um caso inverso. Acreditamos que noruegueses, dinamarqueses e suecos sejam três línguas diferentes, cada uma pertencente a seu respectivo país. Ocorre, no entanto, que os códigos linguísticos usados nos três países são muito parecidos, de tal forma que noruegueses, dinamarqueses e suecos conseguem conversar entre si de maneira natural, cada qual usando a sua própria língua, sem necessidade de tradução ou recurso à escrita. Por que então não dizemos que esses três países falam uma mesma língua? Mais uma vez, a unidade geopolítica e o orgulho nacional desses países são reforçados pela ideia de que cada um deles possui o seu próprio idioma. Ao usarmos critérios políticos, será sempre muito difícil identificar quando uma língua termina e a outra começa. O linguista Max Weinreich ironizou a imprecisão do conceito político e sociocultural de língua com a seguinte definição: "uma língua é um dialeto que possui exército e marinha".

Exemplos da semelhança entre norueguês, dinamarquês e sueco.

Frase: "Meu nome é John"	Em norueguês: "Mitt navn er John"	Em dinamarquês: "Mitt navn er John"	Em sueco: "Mitt namn är John"
-----------------------------	--------------------------------------	--	----------------------------------

Exemplos da diferença entre mandarim e cantonês.

Frase: "Meu nome é John"	Em mandarim: "Wò de míngzì shì John"	Em cantonês: "Ngòh go mèhng giu John"	Na escrita ideogramática comum aos dois idiomas: 我的名字是約翰
-----------------------------	---	--	---

Fora do gerativismo, muitas são as abordagens linguísticas interessadas em investigar as línguas sob a perspectiva da língua-E. Ao optar por esse tipo de abordagem externa à mente, os linguistas investigarão fenômenos como, por exemplo, a dinâmica das relações entre, de um lado, o código linguístico e, de

outro, a história, a ideologia, a estratificação sociocultural, a arte, a educação. Podem estudar, também, questões relacionadas ao ensino formal do código e dos usos da linguagem considerados prestigiados, ou, ainda, podem concentrar-se nos produtos da atividade linguística, como os textos. Conforme se vê, estudos sobre línguas-E dedicam-se a propriedades socioculturais, intersubjetivas e históricas relacionadas à linguagem ou dela derivadas. Tais propriedades podem, a princípio, ser analisadas independentemente de sabermos o que se passa no interior da mente das pessoas e, por isso, as consideramos *externas* ao indivíduo.

É correto dizer que *língua* identificada como língua-E é a abordagem da linguagem humana subjacente nas ciências sociais, dentre as quais figuram várias áreas da linguística, como o estruturalismo, a sociolinguística, o funcionalismo, dentre outras.

Figura 2.2: Uma língua-E é o principal meio de interação sociocultural entre os indivíduos.

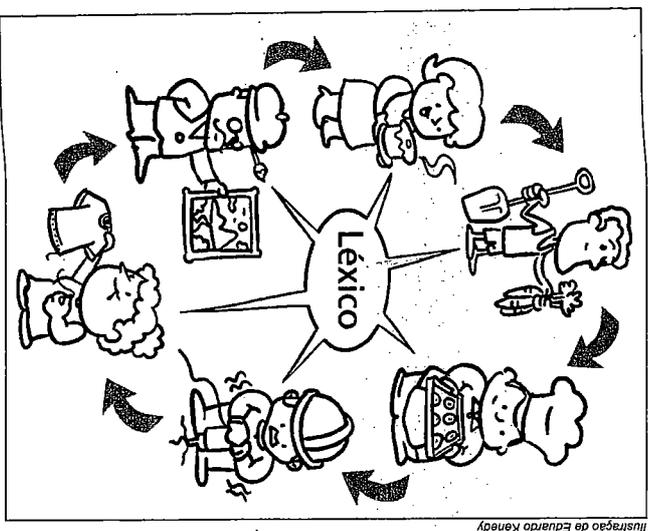


Ilustração de Eduardo Kennedy

Se você já compreendeu o que é a dimensão extensional da linguagem, com o conceito de língua-E, podemos então passar à análise de sua dimensão intensional. Veremos como o código linguístico existente numa língua-E deve ser apropriado por um indivíduo, estocado em sua mente e de lá colocado em uso para a produção de frases e de discursos.

Língua-I

A noção de língua-I corresponde ao conjunto de capacidades e habilidades mentais que fazem com que um indivíduo particular seja capaz de produzir e compreender um número potencialmente infinito de expressões linguísticas na língua de seu ambiente. Dizendo de outra forma, uma língua-I é o conhecimento linguístico de uma pessoa, aquilo que está presente na sua mente e lhe permite usar uma língua-E para produzir e compreender palavras, sintagmas, frases e discursos.

Na acepção de língua-I, uma *língua* é entendida como parte do sistema cognitivo humano. Trata-se de uma faculdade psicológica ou, por assim dizer, um *órgão mental*.

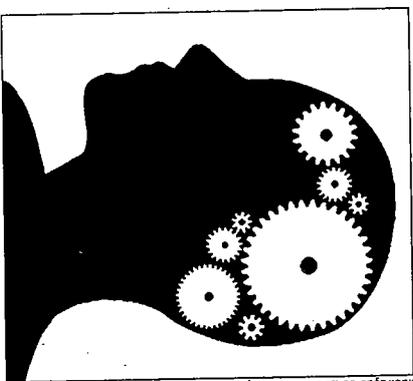
Ao ser concebido, um indivíduo humano herda da evolução de sua espécie um conjunto de instruções genéticas que orientam o desenvolvimento de seu corpo e de seu comportamento. Uma dessas heranças biológicas é a faculdade da linguagem. De posse dessa faculdade, um ser humano normal é capaz de, já ao nascimento, começar a perceber e processar a língua-E de seu ambiente de modo a retirar dela informações para criar a sua língua-I, isto é, o seu conhecimento linguístico. Uma língua-I é a forma pela qual as informações contidas no código linguístico do ambiente (a língua-E) estão representadas em nossa mente. E, digamos assim, a nossa versão interiorizada das informações da língua-E, a nossa cópia particular dessa língua.

Para você entender melhor a distinção entre língua-E e língua-I, pense no seguinte: A língua do ambiente nos oferece fonemas, morfemas, palavras etc. Mas é a nossa língua-I que põe toda essa maquinaria em uso, criando e compreendendo frases e discursos novos a cada momento de uso da linguagem. Por exemplo, usando os códigos da língua portuguesa (seu léxico, sua morfossintaxe etc.), você é capaz de compreender o texto que está lendo neste exato momento. Os códigos que você manipulou para chegar à compreensão do texto são elementos da sua língua-E; porém a habilidade dinâmica de relacionar os itens desse código em frases complexas e extrair deles informação de significado é a sua língua-I, algo que está na sua mente. Da mesma forma, a sua capacidade de usar os códigos da língua portuguesa para produzir e compreender expressões linguísticas quando estiver numa conversa é também manifestação da sua língua-I. Não custa nada reforçar a importante lição da primeira unidade de nosso curso e lembrar a você que nossa habilidade para produzir e compreender a linguagem é inconsciente. Nossa língua-I é, portanto, algo que usamos a todo o momento, de maneira fácil.

No estudo da língua-I, os linguistas estão interessados em descobrir como é a natureza psicológica e neurológica da linguagem na espécie humana. Os estudos dessa abordagem são conduzidos com especial atenção a tudo que é interno à mente das pessoas. Diversos tópicos de pesquisa fazem parte da agenda dos linguistas da língua-I, como, por exemplo, as relações entre linguagem e pensamento, linguagem e memória (de longo e de curto prazo), linguagem e percepção,

linguagem e estrutura neuronal, linguagem e deficiências cerebrais. A abordagem da *língua* como língua-I é, portanto, característica das ciências da cognição — e complementa-se à noção de língua-E típica das ciências sociais.

Figura 2.3: A língua-I é um componente da cognição individual humana.



Pelo conhecimento que, até aqui, você já acumulou desde o início de nosso curso, torna-se fácil entender que o interesse principal da pesquisa da linguística gerativa é a língua-I. O gerativismo é uma abordagem linguística dedicada ao estudo das línguas naturais no sentido de línguas-I.

Caso você já tenha compreendido corretamente a distinção entre língua-E e língua-I, podemos avançar na nossa unidade. Veremos na próxima seção que uma língua-I é, na verdade, um módulo mental, um domínio cognitivo específico. Mas o que é um módulo da mente? É exatamente isso que estudaremos a seguir.

Modularidade da mente

Após compreendermos que o gerativismo se dedica ao estudo da língua-I e, assim, insere-se no campo de estudo da psicologia humana, podemos apresentar outros conceitos fundamentais nas ciências cognitivas que serão muito importantes e úteis no decorrer de nosso curso. Vejamos o conceito de *modularidade da mente*. Uma língua-I é uma capacidade mental, mas existem inúmeras capacidades mentais comuns aos seres humanos. Com efeito, a linguagem é apenas um componente no meio de muitos outros componentes cognitivos existentes na mente. Dizemos que uma língua-I é um *módulo* único e específico na cognição humana. Mas o que devemos entender por módulo?

O conceito de módulo é derivado da hipótese da modularidade da mente. Essa hipótese opõe-se à hipótese da uniformidade da mente. Vejamos como. Segundo a ideia de *mente uniforme*, também chamada pelos psicólogos de *mente holista* ou *inteligência única*, a nossa inteligência seria um todo indivisível. Nessa concepção, acredita-se que a mente humana seja uma espécie de ferramenta única, genérica e de multíuso, a qual utilizaríamos para resolver qualquer tipo de tarefa cognitiva. Para entender melhor a hipótese da uniformidade, imagine que a mente humana seja um computador. Segundo a hipótese holista, esse computador deve possuir somente um único programa, que seria capaz de realizar todas as tarefas cognitivas que dão conta de nosso comportamento normal e cotidiano.

Se precisamos conversar oralmente com alguém, então o programa único está lá fazendo isso com a gente. Se precisamos ler ou escrever uma carta, o programa mental que usaremos é o mesmo. Se precisamos resolver um problema matemático, é o mesmo programa que será utilizado. Se precisamos decidir se temos tempo ou não para atravessar a rua antes que o próximo carro chegue a nossa posição, está lá o programa “pau pra toda obra” fazendo a inferência para a gente. Uma mente uniforme seria, então, essa capacidade única e genérica, que utilizaríamos para todas as formas de comportamento humano.

Figura 2.4: Representação da hipótese da uniformidade da mente.

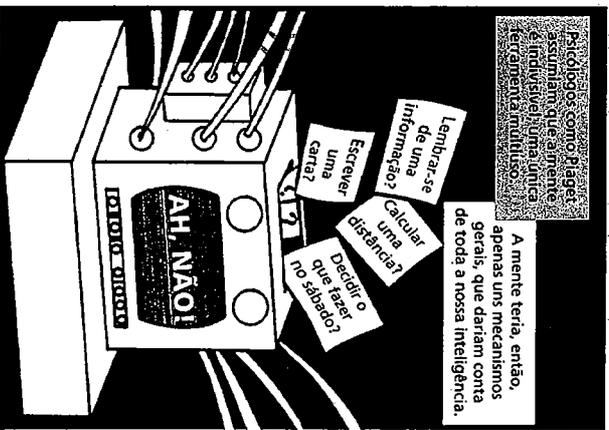


Ilustração de Eduardo Kennedy

A uniformidade da mente encerra, na verdade, uma visão bastante simplificada da cognição humana. No entanto, essa visão orientou o trabalho de filósofos e cientistas durante muitos séculos e, ainda hoje em dia, subsiste. Mesmo o francês Jean Piaget, um psicólogo moderno bastante famoso e influente, assumiu em seus estudos a suposta unicidade da mente: a mente seria a capacidade humana geral e abstrata de aprender por associação, regularidade e analogia. Segundo Piaget, tal capacidade única seria, ela mesma, aplicada a todos os domínios da inteligência, desde a linguagem até a matemática, passando pelas relações sociais, habilidades motoras e técnicas etc.

Com o conceito de modularidade – termo disseminado pelo filósofo e psilinguista norte-americano Jerry Fodor em seu importante trabalho de 1983 –, entendemos que a mente humana seja muito diferente de uma ferramenta única multíuso. Ela é, na verdade constituída, por diversos compartimentos, isto é, divide-se em módulos especializados na execução de tarefas específicas. Isso quer dizer que o que entendemos por *inteligência* é, de fato, um conjunto de *inteligências* especializadas e autônomas, como a linguagem, a visão, a memória, a percepção espacial, as relações lógico-matemáticas etc. A esse conjunto de inteligências (os módulos), chamamos de *mente*, esse termo no singular com apenas cinco letras.

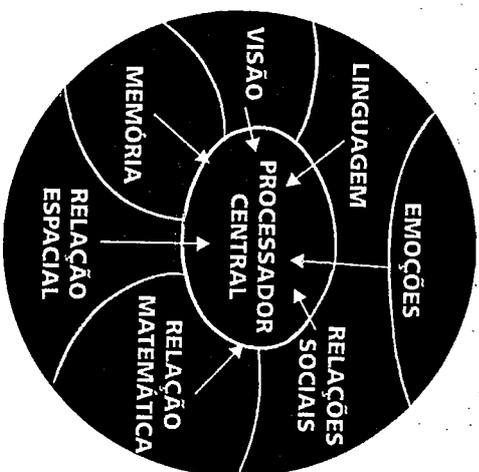
Se voltarmos à nossa comparação entre a mente humana e um computador, diríamos que nesse computador há muitos programas, cada qual especializado num tipo específico de nosso comportamento. Se vamos usar a linguagem, então um programa específico (por exemplo, o Aplicativo Língua-1) é acionado. Se vamos calcular uma distância, para decidir se é possível para atravessar uma rua enquanto o próximo carro não chega perto, então outro programa é acionado (por exemplo, o Aplicativo Movimento no Espaço). Se queremos nos lembrar de alguma coisa, rodamos um novo programa (por exemplo, o Aplicado Busca na Memória) e assim por diante. Uma mente modular é, pelo que dissemos, composta por capacidades especializadas em diferentes tipos de comportamento, dedicadas a diferentes tipos de informação. Cada uma dessas capacidades especializadas é um *módulo* mental.

Conceitos de Fodor

“A mente não pode ser genérica.

Na verdade, ela é um conjunto de inteligências especializadas, cada qual controlada por suas próprias estratégias.”

Figura 2.5: Representação da hipótese da modularidade da mente.



A hipótese de que a mente humana seja modular vem sendo sustentada por evidências cada vez mais contundentes, tanto por parte das ciências cognitivas quanto por parte das neurociências. Por exemplo, sabemos hoje que as pessoas possuem inteligências variáveis, com destaque para certos tipos de inteligências. Isso quer dizer que é normal que, em qualquer pessoa, certos módulos se sobressaíam em relação aos demais. Como ilustração para esse fato, pense que uma pessoa pode comportar-se como excelente orador ou escritor, mas pode não ser tão bom em matemática... Ou pode destacar-se na música, mas não ser tão bom no relacionamento com outras pessoas... E assim por diante. O psicólogo e educador Howard Gardner compreendeu bem a natureza modular da mente e adaptou-a aos problemas escolares em seu livro *Estruturas da mente*. A adaptação de Gardner é muito importante, pois com ela a escola e a sociedade devem tornar-se conscientes de sua responsabilidade em reconhecer e estimular os diferentes tipos de inteligência nas pessoas, sem se concentrar em apenas algumas delas, como infelizmente tem acontecido ao longo da história. Gardner demonstrou, inclusive, que os famosos testes de QI são na verdade uma falácia, afinal eles se baseiam na hipótese da uniformidade da mente e consistem num conjunto de perguntas capciosas que poderiam medir somente a inteligência linguística e a matemática.

Howard Gardner e a teoria das inteligências múltiplas

Gardner vem propondo, em seus estudos, que a inteligência humana seja composta por um complexo de inteligências específicas. Seriam as inteligências visual/espacial, musical, verbal, lógica/matemática, interpessoal, intrapessoal, corporal/cinestésica, naturalista e existencialista. Na verdade, Gardner assume um interesse pedagógico e social no estudo sobre as múltiplas inteligências e não um interesse estritamente cognitivista. Por essa razão, as inteligências por ele propostas não são exatamente módulos, mas alguns módulos combinados, assim como nem todos os módulos da mente são citados nas inteligências elencadas pelo autor.

Também nas ciências do cérebro a hipótese da modularidade da mente vem sendo confirmada. Sabemos hoje que danos em certas áreas do cérebro podem causar deficiência ou perda de algumas habilidades cognitivas. O interessante é que perdas cognitivas decorrentes de problemas como AVC, isquemias ou traumas são quase sempre seletivas, isto é, na maioria dos casos afetam apenas uma parte da cognição, um módulo da mente ou um fragmento desse módulo. Um dano cerebral raramente compromete toda a cognição humana. Certos módulos podem ser afetados, mas muitos são preservados. Um bom exemplo são as afasias, que afetam apenas certas partes do módulo linguístico e do módulo motor, mas não toda a cognição dos indivíduos com o cérebro lesionado. Há também o caso inverso, como acontece na rara síndrome de *savants*, que se dá quando uma pessoa possui um severo retardo mental em muitos módulos da cognição, exceto um ou dois, que curiosamente se apresentam hiperdesenvolvidos, como se compensassem a perda dos outros módulos concentrando em si uma enorme habilidade incomum. Déficits seletivos nas capacidades mentais ou preservação seletiva dessas capacidades indicam justamente que nossa psicologia possui domínios especializados para cada função cognitiva, tal como prevê a hipótese da modularidade da mente.

Se você compreendeu corretamente o que é a modularidade da mente, então será fácil concluir que a linguagem é um de nossos módulos cognitivos. Podemos dizer que a linguagem é um módulo porque ela possui a sua própria natureza e o seu próprio funcionamento, que são independentes dos outros módulos da mente. Nossa inteligência linguística, isto é, nosso módulo da linguagem, não se confunde com a visão, com as emoções, com a inteligência matemática etc. A linguagem humana exerce na mente a função exclusiva de produzir e compreender estruturas linguísticas. Você certamente já entendeu que uma língua-I é o nosso módulo especializado em linguagem, o nosso programa mental para a língua.

A hipótese da modularidade da mente é importante para a linguística porque, com base nela, podemos nos concentrar naquilo que uma língua-I possui de

específico. Os interessados na natureza, na aquisição e no uso de uma língua-1 devem, então, procurar identificar as características próprias desse módulo da mente, ou seja, devem descobrir e descrever as propriedades da mente humana que são exclusivamente linguísticas e não se confundem com a natureza dos demais módulos mentais.

É importante ressaltarmos que o conceito de modularidade da mente não deve ser confundido com a noção de *localizacionismo cerebral* ou de *especificidade neuronal*. Na hipótese localizacionista, assume-se que certas áreas específicas do cérebro e certos tipos de neurônios sejam especializados numa e somente numa atividade neurocognitiva. Dessa forma, um conjunto específico de neurônios numa área particular do córtex cerebral seria especializado numa única função mental, como por exemplo a atenção ou relação espacial. Porém, como muito bem demonstra o neurocientista Miguel Nicolelis, em seu ótimo livro *Muito além do nosso eu* (2011), uma atividade cognitiva em particular pode ser executada por vastas redes de neurônios que se espalham pelo cérebro e trabalham em conjunto. Portanto, você deve entender que a modularidade da mente assume que existam cognições especializadas em certos domínios do saber humano – os módulos –, não obstante, um único módulo cognitivo pode ser executado fisicamente por diversas regiões do cérebro.

Se você já compreendeu o que é a hipótese da modularidade da mente, então certamente compreenderá o que é a modularidade da linguagem. Explicaremos, na próxima seção, como cada módulo da linguagem humana, no interior de nossa competência linguística, possui sua própria estrutura e seu próprio funcionamento.

Modularidade da linguagem

No estudo da mente humana, os psicólogos cognitivos deram um passo à frente e estenderam o conceito de modularidade da mente. Para esses estudiosos, os módulos mentais podem ser, eles mesmos, organizados em outros módulos menores, ou seja, os módulos podem possuir seus módulos internos, também chamados de *submódulos* ou *micromódulos*. Numa comparação grosseira, mas ilustrativa, imagine uma grande laranja. Se você dividi-la em quatro gomos, cada um deles representaria um módulo. Você distribuiria cada módulo desse para uma pessoa diferente, que poderia fazer com ele o que quisesse. Agora imagine que uma pessoa dessas dividisse o seu módulo em várias partes ainda menores, os subgomos da laranja. Cada um desses

subgomos seria o módulo dentro do módulo, ou submódulo. Esses submódulos seriam distribuídos a mais outras pessoas, que poderiam, mais uma vez, fazer o que quiserem com o seu quinhão da laranja. Se você entendeu a comparação, então já compreende também a ideia da modularidade estendida. Cada módulo cognitivo é especializado num tipo particular de tarefa mental e cada um desses módulos possui uma organização interna, caracterizada em submódulos. Esses submódulos, por sua vez, dão conta de uma tarefa específica dentro de seu módulo. Por exemplo, no módulo da visão, possuímos um submódulo para lidar com a visão de objetos estáticos e outro submódulo para lidar com a visão de objetos em movimento. No módulo da memória, há um submódulo para lidar com o armazenamento de informações instantâneas, usadas somente durante os breves momentos de uma conversa, por exemplo, e há outro submódulo para lidar com o armazenamento de informações de maior durabilidade, chamadas de memória de longo prazo, e assim por diante, em todos os módulos.

No que diz respeito ao módulo da linguagem, a ideia da modularidade estendida é útil para explicarmos o fato de que o nosso conhecimento linguístico, a nossa língua-1, é, na verdade, um conjunto de conhecimentos especializados e relativamente independentes entre si. Vejamos a seguir cada um desses submódulos, descrevendo sucintamente a tarefa cognitiva que a eles compete.

O *módulo fonológico* é o submódulo da linguagem especializado nas estruturas sonoras da língua. É esse o tipo de conhecimento que usamos, inconscientemente, quando reconhecemos e produzimos os sons que organizam os morfemas e as palavras. Por exemplo, conseguimos distinguir a forma da palavra *bode* da forma da palavra *pode* porque nosso módulo identifica nessa oposição o traço fonológico de sonorização, que é fisicamente produzido com a vibração de nossas cordas vocais. Em nossa competência fonológica, sabemos que no fonema /b/ tal traço é positivo, isto é, a vibração acontece, mas em /p/ o traço é negativo. O módulo fonológico também nos mantém informados das transformações que determinado som pode sofrer de acordo com o contexto em que esteja inserido na palavra. Por exemplo, sabemos tacitamente que a oposição entre o fonema /s/ e o fonema /ʃ/ se estabelece no começo de palavras, como acontece em, respectivamente, *caixa* e *cheia*, e também entre vogais, como ocorre na distinção entre *assa* e *acha*. (Atenção para não confundir fonemas com letras! O fonema /s/ pode ser representado pelos grafemas “s”, “ç”, “ss” e outros, enquanto /ʃ/ é representado por “x”, “ch” e outros.) Sabemos, da mesma forma, que a oposição entre /s/ e /ʃ/ não se estabelece no final de sílaba ou no final de palavra. É por isso que uma palavra como *mais* pode receber uma pronúncia variável, já que, no fim da sílaba, [s] e [ʃ] não provocam oposição fonológica – lembre-se, por exemplo, da pronúncia

[mais] típica de um paulista e da pronúncia [mai] típica de um carioca. Nosso módulo fonológico controla, portanto, as informações a respeito dos fonemas que existem na língua de nosso ambiente. Essas informações dizem respeito aos traços constitutivos dos fonemas e a como tais traços entram ou não em ação para distinguir formas. Construímos esses conhecimentos durante a infância, nos anos de aquisição da língua, a partir de nossa experiência com a língua-E a que fomos expostos. Depois da infância, colocamos o conhecimento armazenado no módulo fonológico em uso a todo o momento no dia a dia da linguagem.

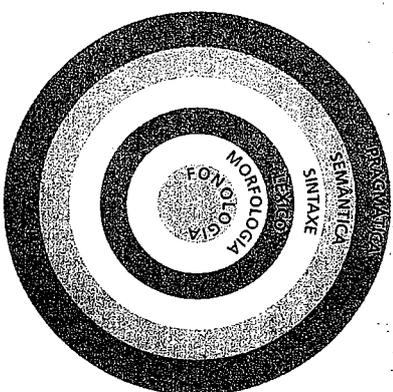
O *módulo morfológico* é o submódulo da competência linguística responsável por nossa habilidade de manipular morfemas. Os morfemas são os constituintes internos da palavra dotados de significação. Eles cumprem as funções de estruturar e modificar os significados no interior dos itens lexicais, bem como nos permitem reconhecer e criar palavras novas. Por exemplo, de posse de um item do léxico como o verbo *estudar*, seu conhecimento tácito sobre a morfologia do português permite a você a expressão de vários significados a partir da adição e/ou subtração de morfemas a tal verbo: {stud-a-re-i}, {stud-a-mos}, {stud-a-va-m} etc. O morfema *-re* adiciona ao verbo o significado de futuro, dando a ideia de que a ação de estudar ainda não aconteceu. Por contraste, o morfema *-va* indica uma ação que ocorreu no passado. Já o morfema *-i* indica que a ação recai sobre a pessoa que fala (*eu* – a primeira do singular), por oposição ao morfema *-mos*, que transmite informação da primeira pessoa do plural (*nós*) e ao morfema *-m*, que aponta para a terceira pessoa do plural (*eles* ou *elas*) realizada o ato de estudar. Além disso, a competência morfológica permite que decodifiquemos o significado de palavras novas, se elas estiverem estruturadas por morfemas. Por exemplo, mesmo que você nunca tenha visto até agora a palavra *desmexicanização*, saberá dizer qual é o seu significado: ato de deixar de ser mexicano. Isso é possível porque o módulo morfológico de sua língua-I faz com que você seja capaz de, inconscientemente, perceber que essa palavra é composta pela relação dos significados expressos pelos morfemas {des-mexic-an-iza-ção}. Quando produzimos e compreendemos modificações de singular e plural, masculino e feminino, passado, presente e futuro, primeira, segunda e terceira pessoas etc., bem como quando produzimos e compreendemos palavras novas, estamos fazendo uso de nossa competência morfológica, que é armazenada no módulo morfológico da língua-I.

Em nossa competência linguística, o *módulo lexical* é responsável pelo armazenamento e pela recuperação do grande número de palavras da língua que falamos, em sua forma de expressão acústica e seu significado. Trata-se de nosso conhecimento sobre o léxico da língua-E. Assim, o fato de você conhecer uma palavra como, digamos, *advogado*, e saber a sua pronúncia habitual (“a-di-vo-ga-du”), o seu significado comum (“tipo de profissão que lida com leis e legislação”) e o seu

contexto de uso (“descrição de uma classe de profissionais” ou “referência a um indivíduo em particular” etc.) é algo que se torna possível por meio do funcionamento normal do conhecimento lexical presente na sua mente. Dar conta das palavras, sua pronúncia, seu significado e suas condições de uso é a tarefa do módulo lexical.

É importante você notar que o módulo lexical não se confunde com um dicionário. Dicionário é um tipo de livro em que são listadas as palavras socialmente reconhecidas numa língua-E. As diferenças são muitas. Por exemplo, é muito comum que um dicionário contenha palavras que não existem na sua língua-I – por acaso, você conhece a palavra *amplexo*? Talvez não, mas saiba que ela faz parte de muitos dicionários do português. É também muito frequente que palavras existentes na sua língua-I não constem nos dicionários – por exemplo, tente encontrar a palavra *tehan* no famoso dicionário Aurélio... Apesar do seu uso há décadas consagrado no Brasil, esse termo ainda não foi dicionarizado. Um dicionário da língua portuguesa pode conter até 400 mil palavras, enquanto a língua-I de uma pessoa muito culta e letrada pode armazenar as informações de 50 a 100 mil palavras. Trata-se, portanto, de duas realidades diferentes. É o módulo lexical que interessa à linguística gerativa. Ele é a parte de nosso conhecimento linguístico especializado nas palavras. É ele que está em formação na nossa fase de aquisição lexical, na infância, e em ação durante nossa vida adulta, quando emitimos e reconhecemos itens lexicais.

Figura 2.6: Representação dos submódulos da linguagem.



É a nossa capacidade de produzir e compreender frases? Ela é levada a cabo pelo *módulo sintático*. Nosso conhecimento tácito sobre a sintaxe é posto em uso cada vez que ouvimos/lemos ou falamos/escrevemos sintagmas e frases. Por exemplo, se quisermos formar uma expressão com as palavras *amigo*, *meu* e *o*,

sabemos que um arranjo possível se dá na forma do sintagma [o meu amigo], enquanto o arranjo *[meu o amigo] é impossível, conforme indica o asterisco. Sabemos disso, inconscientemente, porque nosso módulo sintático controla que tipos de combinação podem ser feitas e quais não podem ser feitas entre palavras, sintagmas e orações. No caso, o artigo *o* pode ser adjunto do sintagma [meu amigo], mas o pronome *meu* não pode ser adjunto do sintagma [o amigo]: [o [meu amigo]] *versus* *[meu [o amigo]].

O submódulo sintático é um dos mais complexos e dinâmicos da linguagem. Você já se deu conta de que, diferentemente do que acontece nos submódulos lexical, fonológico e morfológico, os produtos gerados pela sintaxe são potencialmente infinitos? O número de palavras existentes em nosso conhecimento lexical é, de fato, muito grande e sempre pode ser ampliado, mas, por mais numeroso que seja, é sempre finito. O mesmo se diz acerca de nosso conhecimento sobre a fonologia: são finitos o número de fonemas e o número de combinações possíveis entre os fonemas presentes em nossa competência fonológica. Na morfologia não é diferente. Mesmo que vasto, o número de morfemas que dominamos para criar e modificar palavras é finito. Já o número de sintagmas e frases que podemos construir com o nosso conhecimento sintático é infinito. Com efeito, a cada momento que criamos uma frase, estamos criando uma frase inédita, jamais produzida por nós ou por qualquer outro falante. Você pode confirmar isso por meio de um breve exercício. Escolha aleatoriamente uma frase presente no texto desta unidade e depois vá à biblioteca (ou faça uma busca no Google) e procure encontrar, em qualquer livro (ou site), uma frase exatamente idêntica àquela que você retirou de nosso texto. Ou então, crie uma frase qualquer na sua cabeça e fique esperando até que alguém diga uma frase exatamente igual a essa. Certamente a procura e a espera serão em vão! (Atenção, não estamos falando das frases feitas, aqueles verdadeiros “ditados” que todo mundo repete, como “Água mole em pedra dura tanto bate até que fura”, “Mais vale um pássaro na mão do que dois voando” etc. Estamos falando de frases normais, que usamos cotidianamente para nos comunicar.) Esse aspecto criativo do módulo sintático é uma das propriedades mais importantes da linguagem humana. Falaremos mais sobre ela nas próximas unidades.

O *módulo semântico* tem a função de gerar e identificar significados em expressões linguísticas como palavras, sintagmas e frases. Por exemplo, se perguntassem a você que palavra do português expressa significado oposto ao do item *triste*, você provavelmente indicaria a palavra *feliz* ou *alegre*. Se lhe pedissem para citar uma palavra com o significado relativamente semelhante ao de *longe*, você talvez citasse *distante* ou *afastado*. Além disso, você é capaz de perceber tacitamente a diferença de significado entre a voz ativa e a voz passiva das frases “Cada aluno da minha turma fala pelo menos duas línguas” e “Pelo menos duas

línguas são faladas por cada aluno da minha turma”, ainda que possa ter de pensar com mais cuidado para explicar que diferença é essa. A percepção de relações de significado como as citadas é produzida por nossa competência semântica. Na verdade, o módulo semântico é responsável por muito mais do que isso. É ele que nos torna aptos a produzir e compreender toda a intrincada rede de significação entre os mais diversos tipos de expressões linguísticas, desde os morfemas até a frase.

Com efeito, nem todos os significados produzidos no uso de uma língua natural encontram-se codificados nas palavras ou nas frases. É raríssimo que, no uso cotidiano da linguagem, na interação face a face com os outros falantes, todo o significado do que dizemos seja expresso estritamente pelas palavras e pelas frases que usamos, nada mais e nada menos. Uma grande parte dos significados que veiculamos emana do contexto comunicativo em que nos encontramos. Muito do que significamos tem origem em nossas intenções comunicativas, em nossas pressuposições sobre a pessoa com quem falamos e sobre o assunto de que falamos e em muitos outros fatores que não pertencem à estrutura linguística. Isto é, muitos dos significados da linguagem não são dedutíveis a partir de nosso conhecimento sobre fonemas, morfemas, palavras, frases ou redes de significados. Esse conjunto de fatores, por assim dizer, *extralinguísticos* é o que chamamos de pragmática.

O *módulo pragmático* é, dessa forma, a fração de nosso conhecimento linguístico voltado para o contexto prático de uso da língua. Por exemplo, se chegamos atrasados numa aula e o professor nos pergunta com voz rispida “Você sabe que horas são?”, entendemos, pelo contexto, que isso é uma repreensão pelo atraso e não um pedido de informação sobre as horas. Da mesma forma, se estivermos num jantar com um amigo que precisa ir ao encontro de outra pessoa e ele nos diz “Você sabe que horas são?”, saberemos que isso quer dizer que nosso amigo precisa ir embora logo. Ora, tais significados não estão inseridos na frase citada. Eles se dão no contexto pragmático em que tais frases se inserem. Normalmente, todas as frases que dizemos estão sempre inseridas num dado contexto pragmático, razão pela qual é tão importante que conheçamos os efeitos de sentido que tais contextos produzem.

De fato, a competência pragmática de uma pessoa é um módulo bastante intrínscado e complexo. Ela faz interface com os mais diversos tipos de comportamento social que um indivíduo humano precisa aprender ao longo de sua experiência no mundo real. O conhecimento armazenado em nosso módulo pragmático é, portanto, essencial em nossa tarefa de usar a língua de maneira situada e contextualizada no discurso, caso a caso em nosso dia a dia linguístico.

Bem, após tudo isso que dissemos nesta seção, esperamos que tenha ficado claro para você que, no interior da mente humana, o módulo da linguagem armazena todo o nosso conhecimento linguístico. Esse módulo se distingue dos outros módulos da mente exatamente por ser especializado em apenas um tipo de

função cognitiva: a linguagem. O módulo da linguagem é, na verdade, composto internamente por pelo menos outros seis módulos: fonologia, morfologia, léxico, sintaxe, semântica e pragmática. Como vimos, tais módulos, também chamados de submódulos ou micromódulos, são responsáveis por diferentes tipos de informação linguística especializada. É o conjunto e a interação desses módulos que nos tornam competentes no uso de nossa faculdade linguística, a nossa língua-1. Na seção seguinte, você verá que esses módulos nunca funcionam de maneira separada e isolada. Veremos que, de fato, todos os módulos encontram-se em interação dinâmica durante o uso da linguagem.

A interação dinâmica entre módulos

A modularidade da mente e a modularidade da linguagem não devem passar a você a impressão de que nossa cognição, em geral, e nossa língua, em particular, possuam compartimentos isolados e radicalmente independentes, como se fossem caixas fechadas e separadas entre si por um grande espaço em branco. Nada seria mais falso. Na verdade, todos os módulos e submódulos da mente funcionam de maneira integrada e interdependente, num processo interativo extremamente dinâmico. A descrição dos módulos tal como o fizemos é importante para entendermos as funções que cada um deles desempenha em nossa cognição, todavia a exposição feita não deve jamais dar a entender que exista autonomia absoluta entre os módulos mentais.

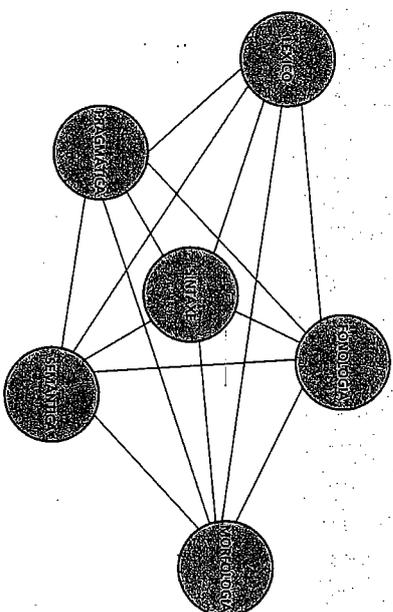
Para tornar evidente a lógica da modularidade e da interação entre os módulos, façamos uma analogia entre a mente e o resto do corpo humano. Quase todos nós conhecemos alguma coisa sobre a *anatomia* de nosso corpo. Com o conceito de anatomia, damos conta do fato de que possuímos diversos *órgãos*. Cada órgão de nosso corpo é justamente um módulo. Assim, cada um deles tem a sua própria estrutura e é especializado numa função particular, por exemplo, o coração bombeia o sangue, o fígado armazena e libera a glicose, o estômago processa os alimentos, os pulmões retiram o oxigênio do ar que respiramos etc. Mas é claro que, no organismo vivo, todos esses órgãos interagem dinamicamente. A essa interação entre os órgãos durante o funcionamento normal do organismo damos o nome de *fisiologia*.

Tal como descrevemos, os módulos da mente são verdadeiros órgãos. A hipótese da modularidade da mente é, portanto, uma espécie de descrição da anatomia da mente. Já a interação dinâmica entre os módulos cognitivos são o que podemos chamar de fisiologia da mente. É exatamente nesse momento em que a mente funciona na prática, na sua fisiologia, que veremos a indelével dinâmica da interação entre

os módulos mentais. Na psicologia cognitiva, tal fisiologia cognitiva é chamada de *processamento mental*. Desse modo, passando aqui a um exemplo bem simplificado, quando nos lembramos de um acontecimento qualquer, usamos o módulo da memória, e imediatamente remetemos essa informação a outro módulo, digamos, o da interação sociocultural, e fazemos isso por intermédio do módulo linguístico, usando uma palavra ou uma frase, por exemplo, sem nos esquecer dos módulos da emoção, que nos põem num estado de ânimo tal durante a interação. Isto quer dizer que, a todo o momento, nossos módulos cognitivos estão ativos e interconectados. Essa é a natureza dinâmica de nossa mente.

O módulo linguístico não é exceção. Usamos fonemas para criar morfemas, que criam palavras, que combinadas geram sintagmas e frases, que expressam certo significado que assume determinado valor pragmático numa situação comunicativa específica no discurso. A cada vez que usamos a linguagem, recrutamos todos os submódulos linguísticos, que estão sempre em ação, um transbordando sobre o outro. O que vimos nesta unidade foram os *órgãos da linguagem*, sua anatomia por assim dizer. Na décima unidade de nosso curso, veremos em mais detalhes a sua fisiologia, isto é, veremos a interação dinâmica entre os módulos.

Figura 2.7: Módulos da mente e submódulos da linguagem estão sempre em intensa interação dinâmica.



Conclusão

Nesta unidade, vimos que o termo *lingua* é ambíguo. Tal ambigüidade, como vimos, pode provocar muitas confusões entre os linguistas e demais estudiosos da linguagem. Para evitar a imprecisão do termo, aprendemos os conceitos de

língua-E e de língua-I. Aprendemos também o conceito de modularidade da mente, por oposição à ideia da mente uniforme, e vimos como esse conceito está relacionado ao de língua-I. Por fim, analisamos os seis submódulos da competência linguística humana e deixamos claro que cada um deles possui estrutura e função específicas no domínio da linguagem, não obstante todos estejam sempre em interação dinâmica e sobrepostos no uso normal que cotidianamente fazemos de nossa capacidade linguística.

Exercícios

- 1) Imagine que você esteja num encontro internacional de estudantes e seja apresentado a Juan, um jovem paraguaio que fala espanhol e guarani. Como você descreveria a dimensão objetiva e a dimensão subjetiva das línguas que o Juan domina?
- 2) Identifique e explique o conceito de *lingua* subjacente na definição abaixo. Uma língua é a instituição pela qual os humanos se comunicam e interagem uns com os outros por meio de símbolos arbitrários orais-auditivos habitualmente utilizados. (Lyons, 1987: 4)
- 3) Analise o excerto abaixo. Identifique e explique o conceito de língua utilizado pelo autor.
A linguagem é uma habilidade complexa e especializada, que se desenvolve espontaneamente na criança, sem qualquer esforço consciente ou instrução formal, que se manifesta sem que se perceba sua lógica subjacente, que é qualitativamente a mesma em todo indivíduo, e que difere de capacidades mais gerais de processamento de informações ou de comportamento inteligente. Por esses motivos, alguns cognitivistas descreveram a linguagem como uma faculdade psicológica, um órgão mental, um sistema neural ou um módulo computacional. Mas prefiro o simples e banal termo *instinto*. (Pinker, 2003: 9-10)
- 4) Considerando a hipótese da modularidade da mente, por que a linguagem deve ser considerada um módulo particular?
- 5) Analise os seguintes casos de pessoas com problemas cognitivos durante a aquisição de uma língua natural. Explique o que tais casos ilustram acerca da modularidade da linguagem.

(a) **O caso Chelsea.**
Chelsea iniciou a aquisição do inglês por volta dos 30 anos de idade. Ela era deficiente auditiva e passou, então, a ser oralizada e alfabetizada. Após anos de contato com o inglês, sua produção linguística apresentava estruturas morfosintáticas anômalas. Vejamos exemplos de frases de Chelsea. *Orange Tim car in* (tradução: "Laranja Tim carro em"), dita para expressar *There is an orange in Tim's car* (tradução: "Há uma laranja no carro de Tim"); *Breakfast eating girl* (tradução: "Café da manhã comendo menina"), dita para expressar *The girl is eating breakfast* (tradução: "A menina está comendo o café da manhã"). Chelsea, no entanto, apresentava uma competência lexical e semântico-pragmática normal.

(b) **O caso Antony.**
Antony foi uma criança acometida de retardamento mental. Seus enunciados eram deficientes do ponto de vista semântico e pragmático. Sua sintaxe, sua morfologia e sua fonologia, porém, eram normalmente elaboradas. Eis uma frase de Antony: *I got two sisters. I got David and Vicki and Margaret* (tradução: "Tenho duas irmãs. Tenho o David, a Vicki e a Margaret").

- 6) Numa crítica comumente feita não só a gerativistas, mas a todos os cientistas que assumem a modularidade da mente, afirma-se que a cognição humana não pode ser, de maneira nenhuma, modular. Afinal, dizem os críticos, quando fazemos uso de uma faculdade mental qualquer, como, por exemplo, a linguagem, lançamos mão de muitos tipos de conhecimento que atuam simultaneamente e, de forma alguma, parecem estar isolados de forma rígida. Explique, de acordo com o que você aprendeu nesta unidade, se esse tipo de crítica representa, de fato, objeção à hipótese da modularidade da mente e à hipótese da modularidade da linguagem.

UNIDADE 8

Sintaxe e computações sintáticas

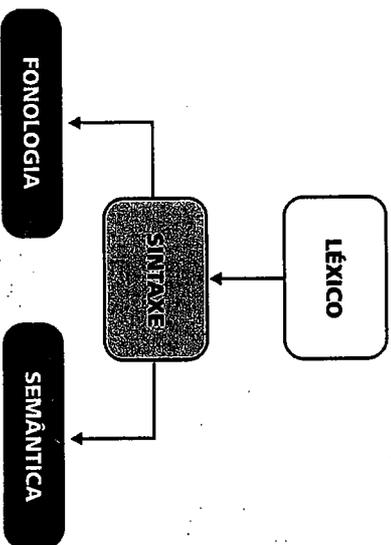
Ao ouvir falar em “sintaxe”, talvez você se recorde das aulas de língua portuguesa nos tempos de sua Educação Básica. Nas escolas brasileiras, “sintaxe” é quase sempre sinônimo de “análise sintática” — aquela tarefa de identificar e classificar a função sintática dos termos da oração e do período (funções como “sujeito”, “objeto direto”, “oração subordinada substantiva completiva nominal” etc.), que, às vezes, tanto inquietava os estudantes. Esperamos que esse tipo de conteúdo gramatical explícito seja familiar para você, mas, caso não seja, não há grande problema. Na verdade, a noção de sintaxe que vamos apresentar e detalhar nesta e na próxima unidade de nosso curso é muito diferente de tudo aquilo que tradicionalmente se ensina na escola. Vamos aprender a identificar na sintaxe o Sistema Computacional da linguagem humana, isto é, pensaremos a sintaxe como o conjunto das operações cognitivas que geram representações sintáticas complexas, como sintagmas e frases. Particularmente, acreditamos que os conteúdos desta unidade ajudarão você a compreender melhor os fundamentos da análise sintática escolar.

Já aprendemos, neste curso, que uma grande porção do conhecimento linguístico humano diz respeito ao Léxico e às computações lexicais. Com efeito, quando somos capazes de produzir e compreender a linguagem, sabemos, dentre outras coisas, evocar e reconhecer palavras, identificar o significado associado a um determinado significante num dado item lexical, sabemos prever certos itens que, numa frase, deverão ser associados a outros e, por fim, conhecemos uma longa e complexa rede de relações semântico-sintáticas entre predadores e argumentos. Não obstante, todo esse vasto conhecimento é apenas uma fração de tudo o que sabemos sobre a língua natural que adquirimos em tenra infância. O grande conjunto da competência linguística humana engloba também a fonologia, a morfologia, a sintaxe, a semântica, a pragmática e o discurso.

O componente da linguagem que mais vem ocupando a atenção da linguística gerativa nos últimos 50 anos é indubitavelmente a *sintaxe*. Essa preferência não é aleatória. Ela justifica-se, de acordo com Chomsky, pelo fato de a sintaxe funcionar como o componente central da cognição linguística humana. Na interpretação

gerativista, a sintaxe cumpre a função de alimentar os sistemas fonológico e semântico com representações linguísticas que, por seu turno, são construídas com base em informações retiradas do léxico e da morfologia. Isso quer dizer que, na visão chomskiana, a sintaxe funciona com uma espécie de “logística central”, a qual estabelece conexões entre todos os demais submódulos da linguagem.

Figura 8.1: A centralidade da sintaxe na arquitetura da linguagem.



A centralidade da sintaxe na cognição linguística torna-se evidente para um principiante nos estudos da linguagem quando ele se dá conta de que os humanos raramente se comunicam por meio de palavras soltas, isto é, itens lexicais isolados e não combinados em expressões complexas como sintagmas e frases. Tipicamente, a interação linguística acontece por meio de discursos complexos, que, por sua vez, se estruturam em frases complexas. Do ponto de vista de quem fala ou escreve, os discursos que compõem a atividade linguística humana cotidiana são constituídos serialmente, passo a passo, por meio das frases que codificam e expressam nossos pensamentos. Para quem ouve ou lê, os discursos também são constituídos de fração em fração, à medida que o sinal linguístico é decodificado nas frases e integrado nos discursos que permitem a troca de estados mentais entre as pessoas.

Ao longo de todo o caminho entre a palavra e o discurso está a sintaxe. “Sintaxe” é, portanto, a fração de nossa cognição linguística que lida com sintagmas e frases. É a nossa habilidade de combinar unidades do léxico para formar e identificar representações complexas de maneira recursiva, gerando as orações e os períodos que compõem os discursos. Descubrir *como é a natureza da sintaxe no interior de nossas mentes* é um dos principais empreendimentos da linguística gerativa.

Com efeito, antes de iniciarmos propriamente o estudo da sintaxe na mente, investigando em mais detalhes o funcionamento do Sistema Computacional da

linguagem humana, há uma série de princípios analíticos e notações terminológicas que precisamos dominar. A presente unidade do curso cumpre justamente o papel de apresentar a você os conceitos e as ferramentas de análise que são fundamentais para o estudo da sintaxe. Prepare-se para uma unidade repleta de tecnicidades, as quais representam os primeiros passos no ensino da metalinguagem necessária ao trabalho do **sintaticista de orientação gerativista**.

Sintaticista é o tipo de linguista que se especializa no estudo da sintaxe das línguas naturais.

É hora de você aguçar sua curiosidade intelectual e manter sua atenção em capacidade máxima, pois estamos começando nossa incursão por uma das regiões mais importantes e mais interessantes do conhecimento linguístico humano: nossa competência sintática.

Sintaxe: o Sistema Computacional

Já dissemos que o termo “sintaxe” possui um valor bem específico para os gerativistas. Ele se refere à nossa capacidade de construir representações linguísticas complexas, como sintagmas e frases. Nessa acepção, sintaxe possui pouca relação com o estudo das funções sintáticas, das regras de concordância e de colocação de termos na frase que fazemos nas aulas tradicionais de língua portuguesa. Sendo assim, a primeira coisa que precisamos estabelecer nesta unidade é que o termo “sintaxe” está aqui sendo usado como sinônimo de Sistema Computacional da linguagem humana.

Sintaxe é uma palavra tradicional que, nas ciências cognitivas, assume o valor de Sistema Computacional da linguagem humana. Trata-se do conjunto das computações cognitivas que geram representações linguísticas complexas, como sintagmas e frases.

Para o gerativismo, a sintaxe é, portanto, um atributo da língua-1 presente na cognição de cada indivíduo particular. Note que, ao falarmos de “sintaxe” como uma propriedade da mente humana, podemos ter a ilusão de que os fenômenos sintáticos são sempre os mesmos em todas as línguas naturais. Na verdade, em termos de universalidade e de particularidade, a sintaxe não é diferente do restante de nossa cognição: certos fenômenos sintáticos são universais e outros são variáveis. Por exemplo, lembre-se da unidade deste curso em que tratamos da

Teoria de Princípios e Parâmetros. Você certamente se recorda de que na sintaxe de todas as línguas naturais existem relações computacionais que se estabelecem entre um determinado núcleo lexical e seu respectivo complemento (argumento interno). A seleção de complemento por um dado núcleo é uma boa ilustração de uma propriedade sintática universal. Por outro lado, a posição linear que um núcleo ocupa em relação a seu complemento ilustra bem uma propriedade variável entre as línguas. Há línguas em que o núcleo antecede o seu respectivo complemento — como é o caso do português —, e há línguas em que o núcleo sucede o seu complemento — como acontece no japonês. Isso quer dizer que, como sintaticistas, temos a função de observar, descrever e explicar os fenômenos sintáticos que são específicos de uma língua ou de um conjunto de línguas e os fenômenos sintáticos que são universais, invariantes no Sistema Computacional.

Os diferentes fenômenos sintáticos que observamos nas línguas naturais são o resultado da interação entre o Léxico e o Sistema Computacional. O Léxico, como já sabemos, é intensamente variável de uma comunidade linguística para a outra, mas o Sistema Computacional é, na verdade, invariável dentro os indivíduos humanos. Isso significa que a variabilidade dos fenômenos sintáticos resulta das informações lexicais que o Sistema Computacional usa para criar representações linguísticas. Ou seja, o Sistema Computacional é sempre o mesmo em todos os indivíduos, enquanto os fenômenos sintáticos são variáveis entre as línguas.

Na condição de Sistema Computacional, a unidade linguística mínima que importa à sintaxe é a *palavra*. Naturalmente, existem muitos tipos de informação linguística inferiores à palavra (tais como o fone, o fonema, a sílaba, o morfema), mas eles não são acessados pelo Sistema Computacional. Isso quer dizer que a entidade mínima capaz de desencadear operações sintáticas é a palavra, o item lexical que carrega um conjunto de traços.

Neste momento, você pode interrograr-se: se o item lexical é a unidade mínima das computações sintáticas, qual deve ser então a sua unidade máxima? A resposta para isso é “a frase”. A sintaxe cumpre, na dinâmica da linguagem humana, a função de combinar unidades até o limite da frase. As computações linguísticas em nível superior à frase já não são executadas pela sintaxe, mas, sim, pelo discurso, numa complexa rede de relações semânticas e pragmáticas inacessível ao Sistema Computacional. Dito isso, você pode facilmente compreender que o trabalho da sintaxe, no interior de nossa cognição, começa com as palavras e termina com a frase — mas não se esqueça de que há um vasto oceano de computações linguísticas àquém da palavra e além da frase.

Pois bem, se a unidade mínima das operações sintáticas é o item lexical e se sua unidade máxima é a frase, qual deve ser, então, a *unidade intermediária* da sintaxe? Para respondemos a essa pergunta, apresentaremos na seção a seguir o conceito de *sintagma*. É certo que já vínhamos utilizando o termo ao longo deste curso, mas o fizemos de uma maneira um tanto informal. Até aqui, entendíamos sintagma como um conjunto de palavras. Sejam agora um pouco mais rigorosos e pensemos: o que são os sintagmas e como eles são relevantes para as computações sintáticas? Vejamos isso em detalhes.

A noção de sintagma

A noção de sintagma é derivada da noção matemática de *conjunto*. Um conjunto é tipicamente uma coleção de unidades que formam um todo, uma unidade complexa. Da mesma forma, um sintagma é tipicamente um conjunto de elementos, uma unidade complexa.

Dizemos que o sintagma é uma *unidade* porque computacionalmente ele se comporta como tal. Um sintagma, ainda que seja constituído internamente por diversos elementos, é manipulado pelo Sistema Computacional como se fosse uma peça única. Nos exemplos a seguir, vemos que, ao deslocar o argumento interno (que aparece destacado entre colchetes) de um verbo para o início da frase, deslocamos na verdade o conjunto de elementos que constituem esse argumento, independente de quantas palavras sejam.

- (1) a. João viu [Maria] na festa.
b. [Maria], João viu na festa.
- (2) a. João comprou [aquele livro] no shopping.
b. [Aquele livro], João comprou no shopping.
- (3) a. O meu primo fez [aqueles complexos exercícios que a professora passou] em menos de uma hora.
b. [Aqueles complexos exercícios que a professora passou], o meu primo fez em menos de uma hora.

No exemplo (1), o argumento interno do verbo “ver” é constituído por uma única palavra. Essa palavra é computada como uma unidade pelo Sistema Computacional e, assim, é deslocada para o início da frase em (1b). Trata-se, portanto, de um sintagma que, no caso, possui apenas um constituinte interno. Já nos exemplos (2) e (3), o argumento interno dos verbos “comprar” e “fazer” é complexo. Ele possui mais de um constituinte. Em (2), o sintagma “aquele livro” é constituído

por duas palavras e, em (3), o sintagma “aqueles complexos exercícios que a professora passou” constitui-se por meio da relação entre diversos elementos: o sintagma “aqueles complexos exercícios” e a oração “que a professora passou”. Nesse último caso, vemos que um sintagma pode inclusive constituir-se de outros sintagmas. O interessante é que o deslocamento de tais sintagmas para o início da sentença, em (2b) e (3b), ocorre tal como o deslocamento de (1b). O Sistema Computacional realizou a operação de deslocamento sobre esses elementos independentemente da constituição interna de cada um deles, como se todos fossem somente uma única unidade, um bloco a ser deslocado na frase: um *sintagma*. Isso significa que um sintagma constitui-se como tal se o Sistema Computacional puder manipulá-lo como uma única unidade, a despeito de sua complexidade interior.

Um sintagma é tipicamente um conjunto de unidades (seja um conjunto de palavras ou de outros sintagmas). Entretanto, um sintagma pode também ser constituído por somente uma palavra ou mesmo por nenhum elemento foneticamente realizado na frase. Para entender isso, lembre-se de que o conceito de sintagma é derivado do conceito de conjunto. Você deve lembrar-se da existência do *conjunto unitário* e do *conjunto vazio*. Para a sintaxe, o conjunto unitário é o sintagma formado por uma única palavra, enquanto o conjunto vazio é formado por um elemento sem matriz fonética. Por exemplo, na pergunta “Você conhece Paulo?”, o argumento interno “Paulo” é um conjunto unitário, um sintagma com um único constituinte. Já se a resposta a essa pergunta fosse “Sim, eu conheço”, veríamos que argumento interno (no caso, o objeto direto) do verbo “conhecer” não possui agora qualquer conteúdo fonético, comportando-se como um conjunto vazio ou, como se diz no gerativismo, uma categoria vazia (“pro”).

Como fazemos para identificar sintagmas numa dada frase? Essa é uma pergunta importante. Com efeito, existem alguns testes de identificação de constituintes que permitem a localização dos sintagmas na sequência de palavras de um enunciado qualquer. Os mais básicos desses testes são *interrogação*, *pronominalização*, *topicalização* e *elipse*.

O teste da *interrogação* é útil para identificarmos sintagmas porque, quando usamos um pronome interrogativo como “quem”, “o que”, “como”, “onde” etc., tal pronome substituirá um sintagma completo. Ele nunca substituirá apenas parte do sintagma, tampouco transbordará sobre outros sintagmas. Por exemplo, nas frases em (b) a seguir, o pronome interrogativo substituiu o sintagma destacado em colchetes nas frases em (a).

- (4) a. O aluno leu [muitos livros].
b. O aluno leu [o quê]?

- (5) a. [Jorge] leu muitos livros.
b. [Quem] leu muitos livros?
(6) a. O rapaz saiu de casa [sem roupas adequadas para o frio].
b. O rapaz saiu de casa [como]?

- (7) a. Todos os meus amigos foram para [alguma praia deserta] a fim de surfar.
b. Todos os meus amigos foram para [onde] a fim de surfar?

Com esses exemplos, podemos ver que um pronome interrogativo é um sintagma equivalente a outro sintagma qualquer, independente de sua extensão. Em (4), o interrogativo “o quê”, substitui um sintagma que também é constituído por dois elementos: “muitos livros”. Essa coincidência de número de constituintes repete-se no exemplo (5), já que o interrogativo “quem” e o argumento externo “Jorge” são sintagmas unitários, mas se perde em (6) e (7), exemplos em que o sintagma substituído pelo interrogativo é formado por diversos elementos.

Você compreendeu o que é o teste da interrogação? Se sim, então vamos aplicá-lo agora recursivamente sobre os grandes sintagmas destacados em (6a). Perceba que eles são complexos, isto é, são formados por sintagmas dentro de sintagmas. Como podemos identificá-los recursivamente por meio do teste da interrogação? Vejamos.

Note que, ao usarmos colchetes para isolar os sintagmas, temos de ter atenção para abri-los e fechá-los corretamente. Para cada colchete aberto, deve haver outro que o feche. Abrimos e fechamos colchetes nas posições sintagmáticas que desejamos destacar. Assim, em [sem [roupas [adequadas [para [o frio]]]], indicamos a existência de cinco sintagmas, que começam na posição em que o colchete é aberto (virado para a direita). Como desejamos indicar que esses sintagmas estão localizados uns dentro dos outros, deixamos para fechar o colchete de cada sintagma somente ao fim do grande sintagma em que todos estão incluídos. É por isso que, somente ao seu fim, encontramos os cinco colchetes finais juntos (virados para a esquerda).

Em (6a), podemos interrogar [sem [o quê]]? No caso, “o quê” substituiria o sintagma [roupas adequadas para o frio]. Uma vez identificado esse longo sintagma, poderíamos seguir adiante e perguntar: [sem [o quê] adequadas para o frio]? Isso nos levaria a identificar [roupas] como um sintagma independente. Por fim, a interrogação [sem roupas adequadas [para [o quê?]]] nos permitiria a identificação de [o frio] como outro sintagma. Recomendamos que você faça um rápido exercício de sintaxe e use o teste da interrogação para identificar os sintagmas existentes em (7a).

O teste da *pronominalização* é idêntico ao teste da interrogação. Ele consiste em usar um pronome pessoal ou demonstrativo em substituição a um sintagma

completo. Por exemplo, os sintagmas destacados em colchetes nos exemplos em (a) a seguir são substituídos por um pronome em (b).

- (8) a. [O filho da Maria] passou no vestibular.
b. [Ele] passou no vestibular.
- (9) a. O guarda fez [um gesto efusivo com as mãos].
b. O guarda fez [isso].

No exemplo (8), vemos que o pronome “ele” substitui todo o sintagma [o filho da Maria]. Aliás, com esse exemplo, você pode constatar que a definição escolar segundo a qual “o pronome é uma palavra que substitui um nome” está apenas parcialmente correta. Na verdade, o pronome substitui um sintagma, e não um nome como palavra isolada. Já no exemplo (9), o demonstrativo “isso” substitui todo o sintagma [um gesto efusivo com as mãos]. Note que todos os sintagmas destacados em (a) são complexos. Portanto, você fará um ótimo exercício ao identificar a estrutura interna completa desses sintagmas, utilizando para tanto os testes da interrogação e da pronominalização.

O teste da *topicalização* consiste em deslocar um constituinte para a periferia à esquerda da frase, ou seja, para o início da frase. “Topicalizar” é transformar em tópico (colocar no início), e “tópico” é sempre a coisa sobre a qual se fala com destaque na frase. É isso o que acontece com os sintagmas destacados em (b) e (c) a seguir.

- (10) a. O rapaz abriu essa porta com uma moeda.
b. [Com uma moeda], o rapaz abriu essa porta.
c. [Essa porta], o rapaz abriu com uma moeda.

Se compararmos (10b) com (10c), veremos que [com uma moeda] e [essa porta] são dois constituintes independentes, isto é, são dois sintagmas separados. Conforme analisamos em (10b), é possível topicalizar [com uma moeda] e deixar *in situ* (isto é, no lugar de origem) o constituinte [essa porta]. Isso quer dizer que [com uma moeda] é um sintagma por si próprio, já que ele pode sofrer uma topicalização. Por sua vez, [essa porta] também é um sintagma em si mesmo, e sabemos disso porque, como vemos ilustrado em (10c), esse constituinte pode ser topicalizado, deixando *in situ* o sintagma [com uma moeda]. A conclusão é que a topicalização é mais um teste útil para identificarmos o limite entre sintagmas: se um conjunto de palavras pode ser deslocado para o início da frase, então esse conjunto é um sintagma. Caso contrário, então não se trata de um sintagma.

Por fim, o teste da *elipse* é também útil para identificarmos sintagmas. Ele consiste em omitir um constituinte numa estrutura coordenada, fazendo com que

tal constituinte tenha de ser inferido pela pessoa com quem falamos. Quando isso acontece, o constituinte elidido é um sintagma. Vejamos um exemplo.

- (11) a. Paulo [leu o livro] na varanda e João, na sala.
b. João [faltou à aula hoje] e José também.

Em (11a), [leu o livro] é omitido, isto é, sofre elipse, logo depois da palavra “João”. Isso quer dizer que esse constituinte é um sintagma. O mesmo acontece em (11b), em que o sintagma [faltou à aula hoje] é elidido logo depois do advérbio “também”.

O interessante desses testes é que pelo menos um deles funcionará quando você precisar identificar os limites entre os sintagmas numa determinada frase. Por exemplo, se você tivesse de analisar sintagmaticamente uma frase ambígua como (12a), poderia usar o teste da pronominalização (ou outro cabível). Nesse caso, tal teste indicará que há duas estruturas sintagmáticas possíveis, conforme veremos.

- (12) a. O juiz julgou o réu inocente.
b. O juiz julgou-o inocente.
c. O juiz julgou-o.

A ambiguidade em (12a) acontece porque não sabemos se um “réu inocente” foi julgado ou se um dado “réu” foi julgado e o veredito desse julgamento foi “inocente”. Em (12b), a pronominalização do sintagma [o réu] indica que [inocente] é um sintagma independente e, portanto, trata-se do julgamento do juiz – não de uma propriedade do réu. Já em (12c), o fato de pronominalizarmos todo o constituinte [o réu inocente] indica que esse é um único sintagma e, assim, “inocente” é uma característica do “réu” – e não sabemos como o juiz o julgou.

A essa altura, você já deve ter aprendido que, embora possa ser também unitário ou vazio, um sintagma é tipicamente o resultado da combinação (1) de uma palavra com outra, (2) de uma palavra com outro sintagma ou (3) de um sintagma com outro. Isso significa duas coisas muito importantes. Vejamos quais.

Em primeiro lugar, a formação de um sintagma sempre acontece através da combinação entre dois constituintes imediatos (por exemplo, [uma palavra] + [outra palavra] ou [uma palavra] + [um sintagma] ou [um sintagma] + [outro sintagma]). Combinações sintagmáticas são, portanto, binárias. Isso quer dizer que elas não acontecem entre três, quatro, cinco itens etc. combinados todos de uma só vez. Note que as combinações sintagmáticas podem ser diversas e gerar sintagmas extensos e complexos, mas elas sempre acontecem passo a passo com dois itens por vez.

Em segundo lugar, as combinações sintagmáticas podem ser recursivas. Isso significa que o resultado de uma combinação (um sintagma) pode ser usado como um novo constituinte inserido numa nova combinação (formando um novo

sintagma) – e assim sucessivamente. Por exemplo, podemos formar um sintagma com as palavras [o] + [livro]: [o livro]. Depois disso, podemos formar um novo sintagma, combinando [o livro] + [didático], obtendo como resultado [o livro [didático]]. Por sua vez, poderíamos fazer a combinação [o livro [didático]] + [vermelho] e ter como resultado o novo sintagma [o livro [didático [vermelho]]]. Esse último sintagma poderia ser usado numa nova combinação, para gerar uma estrutura ainda mais complexa – e assim por diante.

Embora essa natureza binária da formação de sintagmas pareça ser apenas um recurso descritivo e didático, a combinação binária da sintaxe é, na verdade, uma realidade na percepção e na produção linguística humana. Podemos combinar um número muito grande de constituintes numa frase, mas não fazemos isso combinando todos de uma única vez. Pelo contrário, as combinações seguem uma estrutura hierárquica que é constituída na união de cada dois constituintes por vez. Ainda não se sabe por que razão cognitiva isso acontece, mas, ao que indicam os estudos mais recentes da linguística, combinações binárias são o recurso computacional mínimo, básico e necessário para a criação de estruturas sintáticas recursivas.

Se você já compreendeu o quão complexos podem ser os sintagmas que formamos recursivamente por meio de combinações binárias, talvez lhe reste uma pergunta interessante: qual é o limite das combinações sintagmáticas? Isto é, qual seria o limite de extensão de um sintagma?

Na verdade, teoricamente não há limites. É a princípio possível fazer combinações sintagmáticas ao infinito (com um computador que durasse para sempre e ficasse construindo um sintagma gigantesco até o dia do juízo final). No mundo real, entretanto, os sintagmas encontram limitações impostas pela natureza da cognição humana e pelas motivações comunicativas do uso da linguagem. Nossa memória nos impõe que os sintagmas tenham uma certa limitação, de acordo com o que podemos reter e manipular em nossas mentes. Nossas intenções comunicativas impõem que os sintagmas façam referência a algo no mundo real ou imaginário, sobre o qual desejamos falar. Dessa forma, os sintagmas geralmente possuem uma extensão limitada. A extensão máxima de um sintagma é justamente a frase. Dito isso, você deverá indagar-se: qual seria a distinção entre sintagma e frase? Esse é justamente o assunto da próxima seção.

A noção de sentença

Caso se lembre de suas aulas de sintaxe durante os anos da escola básica, talvez se recorde das noções de *frase*, *oração* e *período*. Numa definição bem tradicional,

dizemos que uma frase é um enunciado linguístico completo, isto é, uma unidade linguística de significação comunicativa por si mesma. Note que uma frase pode ser algo muito simples, como um mero “Oi”, ou algo muito mais complexo, como toda extensão do que estamos dizendo agora desde a palavra “note”. Na linguística, a noção intuitiva de *frase* é capturada pelo conceito de *enunciado*.

Com efeito, não é qualquer enunciado (qualquer frase) que interessa ao sintaticista. A ele só importam as frases que são constituídas a partir de algum predicador verbal. Ora, frases com predicador verbal são aquilo que, na escola, aprendemos a chamar de *orações*. Na linguística, a noção de oração é referida pelo conceito de *cláusula*.

Por fim, o conjunto das orações que compõem uma frase é chamado de *período* nas aulas de língua portuguesa. O sintaticista usa o termo *sentença* para representar o conceito tradicional de período. Quando o período é formado por somente uma oração, chamamo-lo de *período simples* – pois temos na sentença somente uma cláusula. Já quando encontramos no período duas ou mais orações, nomeamo-lo *período composto* – ou sentença complexa, composta por mais de uma cláusula.

É importante que você domine esses termos, pois às vezes podemos encontrar livros ou artigos em que uma palavra é usada pela outra e isso pode deixar-nos confusos. Então, vejamos.

Questões de nomenclatura importantes

Frase = enunciado. Qualquer expressão comunicativa, independente de sua estrutura.

Oração = cláusula. Enunciado constituído em torno de um predicador verbal.

Período = sentença. O conjunto de orações existentes na frase.

A questão que devemos analisar neste momento do curso é: qual é a distinção entre [sintagma *versus* sentença] e [sintagma *versus* oração]? Na verdade, essa distinção a rigor não existe. Orações e sentenças são também sintagmas, pois são igualmente unidades constituídas pela combinação de elementos. Não obstante, a especialização dos termos *oração* e *sentença* (ou período, ou frase) justifica-se pela necessidade de indicarmos a grandeza do sintagma que estamos analisando. Senão, vejamos.

(13) João saiu cedo.

Em (13) temos um enunciado (frase). Tal enunciado é constituído pela combinação imediata de dois sintagmas (os quais são internamente complexos, note bem): [João] e [saiu cedo]. Percebemos que esse enunciado é constituído por uma predicação verbal (com o verbo “sair”), o que nos indica que temos aqui uma

oração – e não um enunciado qualquer. Tal oração é a única em todo o enunciado, isso faz com que, nesse exemplo, a oração coincida com a sentença. (13) é, portanto, uma oração e, ao mesmo tempo, é uma sentença.

Imagine agora que a frase seja a seguinte:

(14) Disseram que João saiu cedo.

Nesse caso, [João saiu cedo] comporta-se como uma oração, isto é, um sintagma com uma predicação verbal, a qual é combinada com [que], gerando o sintagma [que [João saiu cedo]]. Esse sintagma é finalmente concatenado com [disseram] para que a sentença em (14) seja formada. Você deve notar que [disseram] é uma predicação verbal e, assim, encerra outra oração no período. Isso significa que a sentença em (14) é formada a partir da combinação entre dois constituintes unidos pela conjunção “que”: a oração [disseram] e a outra oração [João saiu cedo].

Com esses dois exemplos, entendemos que decidir se [João saiu cedo] é um sintagma, uma oração ou uma sentença depende do momento em que nos encontramos durante a computação da frase. Se [João saiu cedo] é o resultado da última computação num enunciado linguístico, teremos então uma sentença (que, no caso, coincide com a oração). É isso o que ocorre em (13). Entretanto, se houver ainda mais computações a serem executadas até o final da frase, então [João saiu cedo] deverá ser analisado como uma oração, passível de combinação com outras na estrutura do período, tal como acontece em (14).

Em resumo, *sentenças* são o total de sintagmas e orações existentes numa frase, enquanto *orações* é um sintagma específico que apresenta predicação verbal. Sentenças e orações são também sintagmas, mas reservamos esse último termo para fazermos referência a unidades menores do que a oração. Talvez você queira discutir mais sobre esse assunto. Saiba que voltaremos a ele mais à frente em nossa unidade, quando estabelecermos as diferenças sintáticas entre sentenças simples e compostas. Por ora, precisamos aprender que os sintagmas devem ser classificados pelo tipo de item lexical que ocupa o seu núcleo. A tipologia dos sintagmas lexicais das línguas naturais é o tema de nossa próxima seção.

Sintagmas lexicais

Os manuais de sintaxe são unânimes em descrever os sintagmas como unidades *endocêntricas*. Com essa palavra, quer-se dizer que os sintagmas são organizados hierarquicamente em torno de seu núcleo. Já sabemos que construímos sintagmas através de combinações binárias recursivas, o que vamos aprender agora é que, dentre

as duas unidades que fazem a computação sintática que gera sintagmas, uma delas é dominante em relação à outra. Essa unidade dominante é o que chamamos de núcleo.

Um núcleo sintagmático lexical é essencialmente uma unidade do léxico, quer se trate de um predicador ou não. Você deve se lembrar do que aprendemos na unidade anterior, acerca dos traços que compõem os itens lexicais. Vimos que as unidades lexicais básicas são nome (N), verbo (V), adjetivo (A) e preposição (P). Pois bem, esses são os tipos de núcleo que **projetam** os sintagmas lexicais existentes nas línguas naturais.

Chamamos de *projeção* a propriedade de um núcleo lexical expandir-se para o *status* de um sintagma, seja pela combinação com outros elementos, seja pela projeção de si mesmo num sintagma unitário.

Um núcleo é, portanto, uma palavra. É em torno dessa palavra que o sintagma irá estruturar-se. Todavia, nem todas as palavras da língua são núcleos capazes de projetar sintagmas lexicais. Somente núcleos lexicais podem fazê-lo – como o próprio termo “lexical” já anuncia.

Como já dissemos, os núcleos lexicais constituem um pequeno subgrupo das classes gramaticais da língua: N, V, A e P. Sendo assim, podemos prever que um núcleo nominal N sempre projetará um sintagma nominal (SN). Por exemplo, o sintagma [ida [ao teatro]] é um SN, já que se trata da projeção de um núcleo N, no caso o item [ida]. Essa projeção contém não apenas o núcleo [ida], mas também o sintagma [ao teatro]. É a combinação do núcleo N [ida] com o sintagma [ao teatro] que projeta o SN representado a seguir.

(15) [_{SN} ida [ao teatro]]

Pensemos agora no sintagma [fiz [dois cursos de linguística]]. No caso, [fiz] é o núcleo verbal V que seleciona o sintagma [dois cursos de linguística] como complemento. Trata-se, portanto, de um sintagma verbal (SV).

(16) [_{SV} fiz [dois cursos de linguística]]

Como você vê, esse é um sintagma relativamente grande, mas, tal como qualquer outro sintagma complexo, ele é formado pela concatenação entre dois elementos: o núcleo V [fiz] e um outro sintagma [dois cursos de linguística]. (Se você está atento aos exemplos, percebeu que [dois cursos de linguística] é também um sintagma complexo, formado por um núcleo e outros sintagmas – trataremos dele mais adiante nesta seção).

Já em [viciado [em estudar]], [viciado] é o adjetivo A que tem o sintagma [em estudar] como seu complemento. É o núcleo A que projeta o sintagma adjetivo (SA) representado a seguir.

(17) [_{SA} viciado [em estudar]]

Você pode notar, por mais uma vez, que esse sintagma também é formado pela concatenação entre duas unidades. No caso, tais unidades são o núcleo A e a unidade complexa [em estudar].

Por fim, no sintagma [sem [muito tempo]], o núcleo lexical é a preposição P, que, junto do seu complemento, projeta o sintagma preposicional (SP) que vemos a seguir.

(18) [_{SP} sem [muito tempo]]

Trata-se da união binária entre, de um lado, a preposição [sem] e, de outro, o sintagma [muito tempo]. Essa concatenação projeta o SP que você pode examinar no exemplo (18).

Você deve ter percebido que os núcleos lexicais normalmente projetam o seu sintagma mediante a combinação com outros constituintes – os quais, por sua vez, podem ser sintagmas projetados por outros núcleos. Ou seja, a partir dos exemplos citados você pôde notar que, quando, um núcleo lexical se expande para o *status* de um sintagma, é comum que o faça junto de outro sintagma, embora também possa fazê-lo sozinho, como um sintagma unitário. Por exemplo, o sintagma [fiz [dois cursos de linguística]] é um SV porque seu núcleo [fiz] é um verbo V. O interessante é que esse sintagma se compõe, além de V, de um outro sintagma: o SN [dois cursos de linguística]. Vemos, a seguir, que esse SV projeta-se a partir da concatenação entre V e SN.

(19) [_{SV} fiz [_{SN} dois cursos de linguística]]

Ora, se um sintagma pode ter dentro de si um outro sintagma, podemos então deduzir que o SN em (19), que já se encontra dentro de um SV, possa ele mesmo conter dentro de si algum outro sintagma. No caso, é justamente isso o que acontece, conforme veremos a seguir.

(20) [_{SV} fiz [_{SN} dois cursos [_{SP} de [_{SN} linguística]]]]

Vemos em (20) que o SV é constituído imediatamente pelo seu núcleo [fiz] e pelo SN [dois cursos de linguística]. Por sua vez, esse SN é constituído pelo seu

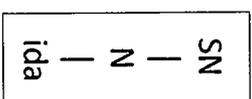
núcleo [cursos], pelo modificador [dois] e pelo SP [de linguística]. Nesse SP, encontramos o núcleo [de], que é uma categoria P combinada com um novo SN, o qual no caso é um sintagma unitário, formado unicamente pelo núcleo N [linguística].

Se esta é a primeira vez que você faz uma análise sintagmática, a identificação de sintagmas dentro de sintagmas pode parecer-lhe um pouco complicada. Nesse caso, as aparências enganam. Com a prática e o treinamento do seu “olho de sintaticista”, rapidamente você será capaz de identificar os núcleos lexicais (N, V, A e P) e as categorias linguísticas que, junto deles, projetam sintagmas. Nessa tarefa de “treinar o olho do sintaticista”, existe um recurso didático muito útil e produtivo entre os gerativistas. Trata-se das representações arbóreas. Tais representações são um recurso visual que podemos usar no lugar da indicação dos colchetes, que vínhamos utilizando até agora. Veremos, na seção a seguir, como é que podemos lançar mão desse recurso descritivo para aprimorar nossa capacidade de identificar e descrever as estruturas sintagmáticas de uma língua natural.

Representações arbóreas

Um diagrama arbóreo – que podemos também chamar simplesmente de *árvore* ou *árvore sintática* – é a representação visual da estrutura de um sintagma ou de uma frase. Numa árvore, as projeções de um núcleo sintagmático são representadas como nós (ramos) e também por linhas (galhos) que indicam os elementos que são concatenados entre si para formar uma projeção. Para exemplificar, pense num sintagma unitário, que é a estrutura sintagmática mais simples possível. Esse tipo de sintagma é formado pela projeção somente de seu próprio núcleo, como ilustrado a seguir.

Figura 8.2: Um SN simples, formado apenas pela projeção de seu núcleo N.



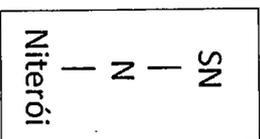
Analisando esta árvore de baixo para cima, podemos notar que sobre o item [ida] há um galho que indica a ramificação do nó N. Esse N é a represen-

tação do tipo de núcleo lexical que temos no sintagma – no caso, o núcleo é N porque [ida] é uma categoria nominal. Como no sintagma não há mais nenhum constituinte além do próprio núcleo, então N projeta imediatamente o seu sintagma SN. Sabemos isso porque, sobre N, há um galho que indica a ramificação do SN, a projeção máxima do núcleo [ida].

Analisemos agora um sintagma mais complexo. E se o sintagma a descrever fosse [ida para Niterói], como seria a sua representação arbórea? Antes de fazermos a descrição desse SN, preste atenção à seguinte orientação. Quando começamos a descrever a estrutura de sintagmas complexos usando representações arbóreas, é muito importante começar nossa análise da direita para a esquerda, desenhando a árvore de baixo para cima. Isto é, ao representar sintagmas numa árvore sintática, devemos começar da última palavra do sintagma e seguir, de trás para frente, até a primeira palavra. Além disso, começamos sempre dos nós mais baixos na estrutura da árvore (os núcleos) e depois subimos os galhos para as representações mais altas, com as projeções dos núcleos. Esse procedimento não é obrigatório, mas pode ser de bastante valia quando estamos desenhando nossas primeiras árvores sintáticas.

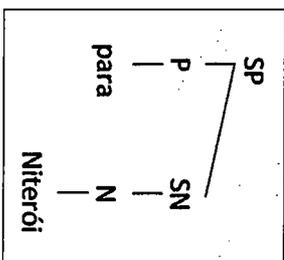
Começamos a representação do sintagma [ida para Niterói] com a projeção do núcleo nominal [Niterói]. Nesse caso, coincidentemente, temos uma estrutura sintagmática idêntica à da representada na figura 8.2., pois mais uma vez encontramos um sintagma unitário, cujo núcleo é o seu único constituinte.

Figura 8.3: Um SN simples que, no caso, é o início da representação de um sintagma complexo.



Na estrutura de [ida para Niterói], o SN representado na figura 8.3 é selecionado como complemento (ou seja, argumento interno) do núcleo preposicional [para]. Com base nessa seleção, teremos uma combinação binária: o núcleo P será concatenado ao SN. É justamente dessa combinação binária que resulta o SP [para Niterói]. Trata-se de um SP porque o núcleo desse sintagma é a preposição [para], uma categoria P, tal como podemos ver na figura que se segue.

Figura 8.4: A concatenação entre P e SN projeta o SP [para Niterói].

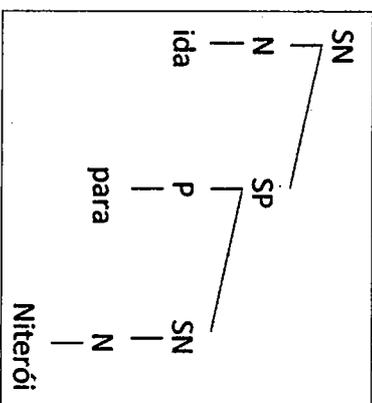


Ao analisarmos essa árvore, podemos rapidamente ver que o SN é projetado diretamente pelo núcleo N, já que não há mais nenhum outro constituinte dentro de seu domínio. Por sua vez, vemos também que o SP é formado pela combinação entre dois constituintes: P e SN.

Em sintaxe, usamos o termo *domínio* para fazer referência a todos os constituintes que estão dentro da estrutura de um dado sintagma. Por exemplo, vemos que, na figura 8.4, o SN está no domínio do SP, mas o SP não está no domínio do SN, ou seja, o SN é desenhado dentro do SP, mas o SP está desenhado fora do SN.

Se você entendeu de que maneira o SP [para Niterói] é formado, então será fácil deduzir qual será a representação do sintagma [ida para Niterói]. Sim, como já sabemos o que é *recursividade*, podemos prever que o processo a formar esse sintagma será idêntico ao destacado na figura 8.4. No caso, o núcleo nominal [ida] será combinado com o SP [para Niterói] e, dessa concatenação, resultará o SN complexo [ida para Niterói].

Figura 8.5: A concatenação entre N e SP projeta o SN [ida para Niterói].



Recapitulando: ao analisar essa figura, você pode perceber que, na representação final do sintagma tomado para exemplo, o N [Niterói] projeta sozinho o SN mais baixo na árvore sintática. Uma vez formado, tal SN é combinado com o núcleo P [para] e, dessa combinação, resulta o SP [para Niterói]. Por fim, esse SP é concatenado com o N [ida], de modo que o SN [ida para Niterói] é projetado. Após essa última computação, nosso sintagma lexical está completo, conforme representamos visualmente na árvore da figura 8.5.

Acreditamos que você já tenha compreendido quais são os fundamentos das representações arbóreas tão usadas pelos sintaticistas: a combinação binária entre elementos, sendo um deles um núcleo. Todavia, antes de apresentarmos uma descrição mais geral e abstrata sobre como representamos qualquer sintagma numa árvore, façamos mais uma análise de um sintagma específico. Dessa vez, vamos analisar um sintagma um pouco mais complexo: [fiz cursos de linguística]. Descreveremos a seguir, passo a passo, como é que esse sintagma é representado numa árvore sintática. Você deve seguir essas orientações e ir desenhado a árvore sintática conforme o que se diz.

- **Primeiro passo.** O primeiro item a representar é o núcleo [linguística]. Trata-se de um núcleo nominal que, portanto, desencadeia o galho que leva à ramificação N. Tal N não é concatenado com nenhum outro constituinte. Isso quer dizer que ele projetará diretamente o SN, que será um sintagma unitário.
- **Segundo passo.** Uma vez projetado o SN [linguística], o próximo elemento que encontramos é a preposição [de]. Ela deve projetar o núcleo P. Esse P tomará como complemento o SN [linguística] formado no primeiro passo. Faremos, portanto, uma combinação binária entre P e SN. É dessa concatenação que deverão subir dois galhos em direção à projeção do SP, um sobre P e outro sobre SN. Note que o sintagma projetado é um SP em função de seu núcleo ser P.
- **Tercer passo.** Já formamos o sintagma [de linguística]. Agora, seguindo para esquerda, encontramos o item [cursos]. Trata-se de um núcleo nominal. Portanto, ele projeta N. Esse N será concatenado ao SP [de linguística] formado no segundo passo. Tal concatenação binária faz surgir a projeção do SN e, assim, chegamos à estrutura [cursos de linguística].
- **Quarto passo.** Finalmente, encontramos o verbo [fiz], que projeta a categoria V. Esse V será combinado com o SN [cursos de linguística], formado tal como descrevemos no terceiro passo. A combinação entre V e SN deve ser indicada pela junção de dois galhos, ligados sobre esses dois sintagmas. É dessa concatenação que se projeta o SV [fiz cursos de linguística].

Se você seguiu corretamente essas instruções, então sua árvore sintática deve ter ficado igual à que representamos na figura 8.6.

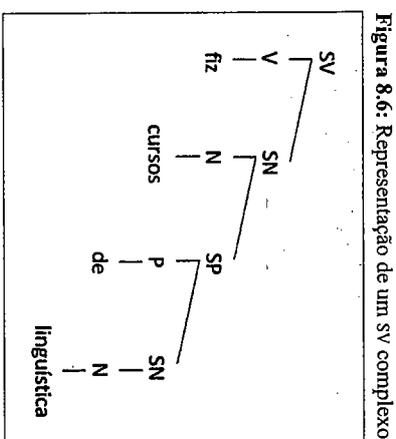


Figura 8.6: Representação de um SV complexo.

As árvores sintáticas são uma ferramenta descritiva bastante útil não apenas para sintaticistas, mas também para gerativistas em geral, psicolinguistas e neuro-linguistas. Se você continuar seus estudos numa dessas áreas da linguística, então saber fazer e interpretar árvores lhe será muito importante. Pratique!

Note bem: árvores sintáticas são apenas um recurso descritivo, uma representação visual útil e didática. Elas não devem ser interpretadas literalmente como o tipo de representação mental que os humanos fazem em tempo real, quando produzem ou compreendem a linguagem.

Teoria X-barra

A lógica das representações arbóreas foi primeiramente formulada por Chomsky, em 1970, e depois foi aperfeiçoada pelo importante linguista Ray Jackendoff, em 1977. Tal lógica é conhecida como *teoria X-barra*. Desde então, as famosas árvores gerativistas vêm sendo vastamente utilizadas como recurso visual para a representação de estruturas sintáticas (e também morfológicas e fonológicas, com adaptações).

Você deve ter percebido que todas as relações entre os constituintes dos sintagmas que até aqui apresentamos como exemplo estabelecem-se entre um dado núcleo lexical e seu respectivo argumento interno. Contudo, sabemos que nem todas as estruturas sintáticas são estabelecidas entre núcleo e complemento. Existem também as relações que se estabelecem entre núcleo e especificador (argumento externo) e, além disso, há ainda aquelas que se dão diretamente entre sintagmas,

como é o caso dos adjuntos. Como esses outros tipos de relação estruturais são representados numa árvore sintática?

A teoria X-barra é interessante e útil justamente porque ela nos oferece um modelo de representação arbórea capaz de dar conta de todos os tipos de relação sintática, seja a de núcleo e complemento, ou a de especificador e núcleo ou ainda a adição entre sintagmas. Vejamos.

Quando o núcleo lexical concatena-se apenas com seu argumento interno ou apenas com o seu argumento externo, a representação sintática que fazemos numa árvore sintática é a seguinte (ver figura 8.7). Note que, na teoria X-barra, usamos a letra X para representar qualquer tipo de núcleo (e o valor de X pode ser qualquer categoria, como N, V, A ou P).

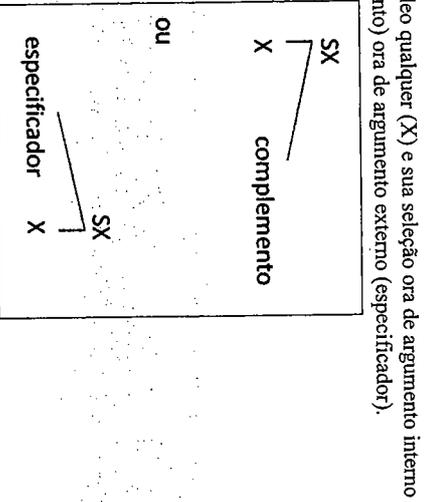
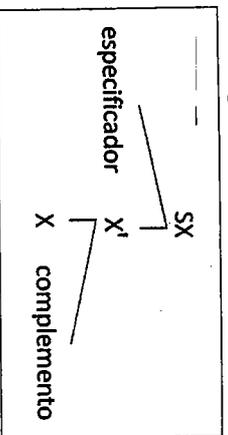


Figura 8.7: Um núcleo qualquer (X) e sua seleção ora de argumento interno (complemento) ora de argumento externo (especificador).

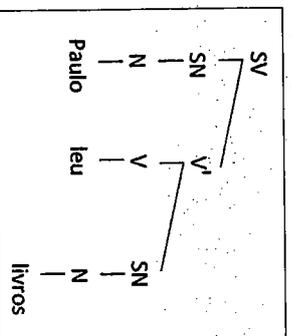
Vemos aqui que um núcleo qualquer (X) será concatenado com um constituinte à direita se este for o seu argumento interno, ou será concatenado com um constituinte à esquerda se este for o seu argumento externo. Ora, o que deve acontecer quando a relação de um núcleo for complexa e, por conseguinte, ele relacionar-se com dois argumentos, um interno e outro externo? Isto é, como devemos representar as duas relações de um núcleo que possui tanto argumento interno (representado à direita) quanto argumento externo (representado à esquerda)? Nesse caso, duplicaremos a representação do núcleo lexical numa nova projeção, que chamaremos de projeção intermediária. Essa projeção intermediária é indicada pela replicação da categoria do núcleo (N, V, A ou P) e pela utilização de uma barra logo ao lado do núcleo: N', V' ou P' – deve-se ler N-barra, V-barra, A-barra e P-barra. É por isso que esse tipo de representação denomina-se X-barra ou X'.

Figura 8.8: Um núcleo qualquer (X) e sua seleção de argumento interno (complemento) e de argumento externo (especificador). X' é a projeção intermediária do núcleo X, e SX é o sintagma do núcleo X, a sua projeção máxima.



A figura 8.8 informa-nos que o núcleo X estabelece duas relações sintáticas. A projeção intermediária X' indica-nos a relação do núcleo com seu argumento interno (complemento) e a projeção máxima SX indica-nos a relação do núcleo com seu argumento externo (especificador). Podemos exemplificar mais concretamente as relações dessa figura com o seguinte sintagma: [Paulo leu livros]. No caso, o predicador [ler] seleciona argumento interno ("livros") e argumento externo ("Paulo"). A representação arbórea desse sintagma deverá ser a que se segue.

Figura 8.9: O predicador [ler], seu argumento interno e seu argumento externo

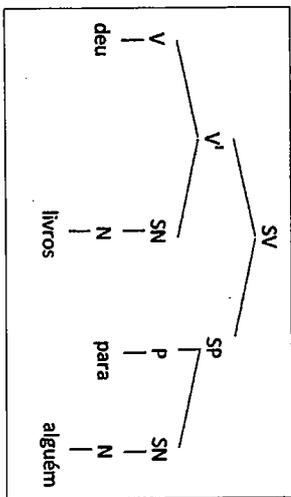


Nessa representação, vemos que o núcleo V possui duas relações sintáticas. É em função dessa dupla relação que utilizamos um nível barra (V'). Primeiramente, o núcleo V seleciona seu argumento interno SN ("livros") e com ele projeta o nível V'. Dada essa projeção, o composto do núcleo mais o argumento interno é concatenado ao argumento externo SN ("Paulo"). É dessa concatenação que se projeta o SV, a projeção máxima do núcleo V.

A duplicação da projeção de X, com o nível X', também acontece no caso dos predicadores que possuem dois argumentos internos, como os chamados verbos bitransitivos. Nesses casos, o núcleo verbal V selecionará, imediatamente, o seu primeiro argumento interno, projetando assim o nível V'. Logo após essa proje-

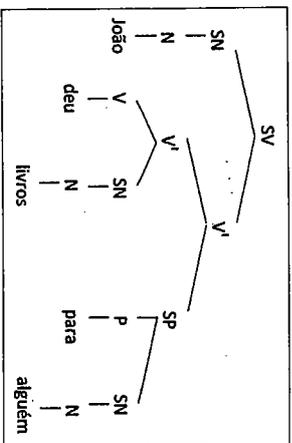
ção, o nível V' será então concatenado ao segundo argumento interno. É isso o que vemos ilustrado na figura a seguir, em que o SV [deu livros para alguém] é composto por duas projeções do núcleo [dar].

Figura 8.10: O nível barra projeto com predadores que selecionam dois argumentos internos.



Há mais uma coisa importante a reter sobre a projeção de um nível barra: ela pode acontecer mais de uma vez na estrutura de um sintagma. Por exemplo, imagine que o SV representado na figura 8.10 tivesse ainda mais uma operação computacional, a qual concatenasse o argumento externo do verbo [dar] à estrutura [deu livros para alguém]. Nesse caso, um argumento externo como "João" seria combinado com uma segunda projeção V' para finalmente dar à projeção máxima de V, conforme vemos a seguir.

Figura 8.11: O nível barra pode ser projetado mais de uma vez numa estrutura sintagmática.



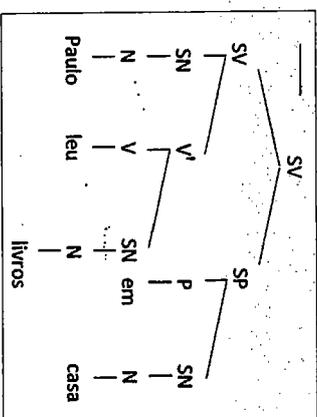
Dessa representação, você pode concluir que o núcleo [dar] concatena-se ao SN [livros], projetando assim o primeiro nível V'. Essa projeção é então combinada com o SP [para alguém], que dá origem ao segundo V'. Finalmente, esse segundo V' concatena-se ao argumento externo, o SN [João], e dessa combinação resulta a projeção máxima SV. Pois bem, já sabemos como um núcleo pode estabelecer uma relação simples, concatenando-se apenas com seu argumento interno ou apenas com seu argumento

externo, e sabemos também como um núcleo pode estabelecer relações complexas, concatenando-se com ambos argumentos interno e externo. Agora vejamos como a concatenação acontece não entre um núcleo e seu(s) argumento(s), mas sim entre sintagmas. Esse tipo de combinação acontece quando temos o fenômeno sintático da adjunção.

Ao analisarmos um sintagma qualquer que se adjunge a outro, devemos ter em conta que as relações estruturais internas a cada um desses sintagmas são estabelecidas normalmente, conforme tudo o que já estudamos até aqui. Isso quer dizer que se adjungirmos, por exemplo, um sintagma como [em casa] a outro sintagma como [Paulo leu livros], a estrutura interna de cada um desses dois sintagmas é dada conforme o que prevê a teoria X-barra. A novidade é a forma pela qual re-presentamos um sintagma como adjunto de outro sintagma. Note que [em casa] é um SP, pois seu núcleo é a preposição [em]. Tal sintagma não pertence à estrutura argumental do predicador [ler]. Trata-se, portanto, de um adjunto.

Na teoria X-barra, adjuntos devem ser representados fora do domínio do sintagma em que se dão as relações entre um núcleo e seus respectivos argumentos. Isso acontece porque a adjunção é uma relação sintática que se estabelece entre duas projeções máximas. É justamente isso o que está representado a seguir.

Figura 8.12: Adjunção do sv [Paulo leu livros].



Nessa árvore, podemos visualizar as concatenações internas do sv [Paulo leu livros], que já descrevemos na figura 8.9, bem como vemos as concatenações internas do SP [em casa]. A adjunção acontece por meio da simples combinação entre esses dois sintagmas: [Paulo leu livros] + [em casa]. Note que dessa adjunção resulta um novo SV. Isso significa que o sintagma que recebe a adjunção e, dessa forma, projeta-se novamente na estrutura sintática é o sintagma verbal. Por sua vez, o SP é o sintagma que se adjunge ao SV.

Se você já dominou a lógica das representações arbóreas, podemos seguir para a próxima seção de nossa unidade. Nesta veremos que, na estruturação das frases

de uma língua natural, outros tipos de sintagma são combinados com os sintagmas lexicais. Trata-se dos sintagmas funcionais, muitas vezes também chamados de sintagmas gramaticais.

Sintagmas funcionais

Já sabemos que SN, SV, SA e SP são sintagmas lexicais, uma vez que se caracterizam como projeções máximas dos núcleos lexicais N, V, A e P, respectivamente. Vejamos agora os sintagmas que são nucleados por categorias funcionais, isto é, palavras que possuem valor puramente gramatical.

A distinção entre categorias lexicais e categorias funcionais é clássica na linguística. Ela foi primeiramente formulada por estruturalistas, com o objetivo de diferenciar, de um lado, palavras de conteúdo lexical, que possuem valor referencial (isto é, remetem-se a algo no mundo biossocial, como “casa”, “amor”, “correr”, “feio”, “alto”), e, de outro lado, palavras de conteúdo gramatical, que possuem valor funcional (isto é, dão conta do funcionamento do sistema linguístico, como artigos, conjunções, numerais, certos pronomes, flexões). Na linguística gerativa, a distinção entre itens lexicais e itens funcionais se estabelece, também, em virtude de suas propriedades selecionais. Uma categoria é dita *lexical* se possui a capacidade de selecionar semanticamente argumentos (s-seleção). Já se tal categoria seleciona argumentos apenas categorialmente (c-seleção), então se trata de categoria *funcional*. Vejamos isso em exemplos.

Uma categoria funcional existente nas línguas naturais é o *complementador* (também chamado de *complementizador*). Trata-se da categoria C que projetará um sintagma complementador (SC). É esse sintagma que marca a força ilocucionária de uma sentença, determinando se se trata de uma frase declarativa ou interrogativa. Além disso, é esse sintagma que estabelece a relação entre orações numa dada sentença. Por exemplo, quando temos numa frase uma oração subordinada e uma oração principal, é o SC que estabelece a relação entre uma oração e a outra, tal como ilustrado a seguir.

(21) [Paula disse [_{SC} que [João está feliz]]].

Note que a oração subordinada [João está feliz] é integrada à oração principal [Paula disse] por intermédio do complementador [que]. Tal complementador, por ser uma categoria C, dá origem à projeção máxima SC. Esse tipo de sintagma não impõe nenhum tipo de restrição semântica aos constituintes que podem ser selecionados como seus argumentos – ou seja, orações com qualquer tipo de conteúdo semântico podem ser introduzidas por um SC. Tal fato nos leva a caracterizar esse sintagma como uma categoria funcional.

Outro sintagma funcional existente nas línguas humanas é o SR, sigla para o sintagma *temporal* (às vezes denominado também de sintagma *flexional*, SF – conforme vimos na unidade 7). Esse sintagma é particularmente interessante porque é ele que atribui a um SV com argumentos saturados o *status* de sentença, conferindo-lhe uma flexão, isto é, uma forma finita (com expressão de tempo, modo, aspecto, número e pessoa em português).

O SR é um sintagma funcional porque ele sempre seleciona um SV como seu complemento, independentemente dos traços semânticos desse SV. De fato, o SR é uma espécie de camada superior à camada lexical de um SV. Vejamos o que isso significa.

Um verbo como [ler] possui a capacidade de selecionar argumentos e atribuir-lhes papel temático, mas não possui em si mesmo a capacidade de marcar a expressão do tempo (ou modo, aspecto, número e pessoa gramaticais). Quem faz esse tipo de marcação gramatical é o SR, que atribui ao verbo um morfema flexional. Isso quer dizer que uma forma de palavra como “lemos”, ou qualquer forma do verbo [ler] ou de qualquer outro verbo, é, na verdade, o resultado da combinação das categorias: a categoria lexical V (a raiz do verbo) e a categoria funcional T (a flexão verbal). Vemos isso representado muito simplificada e a seguir. Somente na próxima unidade veremos a natureza do SR em detalhes.

(22) [Paulo disse [_{SC} que [_{SR} vai [_{SV} ler um livro]]].

Nessa representação, vemos que o SC estabelece a complementação entre a oração subordinada e a oração principal. Logo após esse sintagma, podemos localizar o SR, com a flexão “vai”, que marca o tempo futuro do presente, modo indicativo, na terceira pessoa do singular no SV [ler [um livro]].

Por fim, o último sintagma flexional que devemos considerar é o *determinante*. Um sintagma determinante (SD) é a camada funcional do sintagma lexical SN. É nesse sintagma que se localizam artigos, pronomes, numerais e demais categorias que delimitam informações gramaticais acerca de um nome, tais como definitude (isto é, se trata de um expressão definida ou indefinida), gênero, número, quantidade etc. Por exemplo, nas representações que se seguem, vemos SNs que são, cada qual, determinados de uma maneira diferente por um núcleo D, que projetará um SD.

(23) a. [_{SD} um [_{SN} livro]]
 b. [_{SD} o [_{SN} livro]]
 c. [_{SD} dois [_{SN} livros]]
 d. [_{SD} este [_{SN} livro]]

É importante atentar para a semelhança na função gramatical de SR e de SD. Ambos selecionam um sintagma lexical como complemento, delimitando nele

informações linguísticas (tais como tempo, flexão, definitude, gênero). Tal seleção possui uma natureza puramente gramatical, isto é, não há restrições ao tipo semântico do argumento selecionado – logo, as categorias em questão (T e D) são funcionais.

O complemento de um ST é sempre um SV, enquanto o complemento de um SD é sempre um SN. É em função dessa relação entre sintagmas funcionais e sintagmas lexicais que, normalmente, os SVs são realizados nas línguas naturais em alguma forma verbal específica (num dado tempo, com algum modo e aspecto, em alguma pessoa e número gramaticais) e os SNS realizam-se por meio de uma referência concreta (são definidos ou indefinidos, apresentam uma quantidade específica, pertencem ao masculino ou feminino, estão no singular ou plural).

Ao descrevermos os sintagmas lexicais e funcionais, utilizamos aqui a sigla em língua portuguesa da respectiva abreviatura do sintagma: SN, SV, SA, SP, SC, ST (ou SF) e SD. Ao consultar bibliografia relevante sobre sintaxe gerativa, pode ser que você encontre essas abreviaturas conforme o original em inglês. É o que acontece com o importante livro *Novo manual de sintaxe* (Milošević, Silva e Lopes, 2013). Nesse caso, você deve ter em mente que, em inglês, o termo "sintagma" é traduzido com a palavra *phrase*. Sendo assim, as traduções equivalentes serão as seguintes.

SN = NP (<i>noun phrase</i>)
SV = VP (<i>verb phrase</i>)
SA = AP (<i>adjectival phrase</i>)
SP = PP (<i>prepositional phrase</i>)
SC = CP (<i>complementizer phrase</i>)
ST = TP (<i>temporal phrase</i>)
SF = IP (<i>inflectional phrase</i>)
SD = DP (<i>determiner phrase</i>)

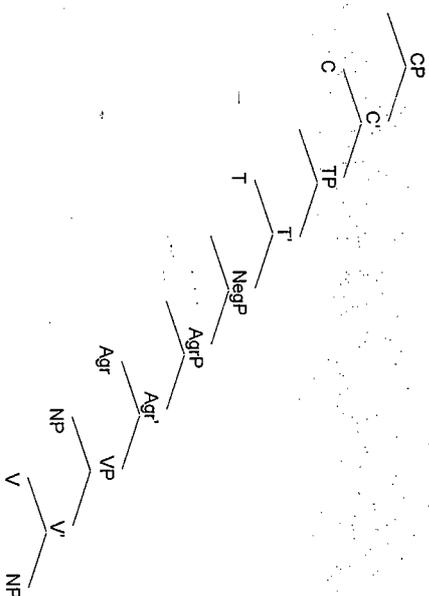
Note que, em inglês, a abreviatura dos sintagmas sempre termina na letra "P", de *phrase* (tradução de sintagma, e não frase – não confundal). É por essa razão que, jocosamente, as teoricidades dos sintaticistas da orientação gerativista são chamadas de "língua do pé".

Na verdade, existem outros sintagmas funcionais importantes nas línguas naturais. Contudo, o estudo mais aprofundado das categorias gramaticais deve acontecer noutro momento, num estágio mais avançado da sua formação acadêmica. Por ora, você deve ter em mente que SC, ST e SD são os principais sintagmas funcionais computados pela linguagem humana. Quando, na próxima unidade, analisarmos as computações sintáticas que realizam a derivação de representações linguísticas, veremos que existem muitas relações estruturais entre categorias lexicais e funcionais. Portanto, fixe esses sintagmas em sua memória e em seu conhecimento.

Para finalizarmos esta unidade, resta-nos descrever rapidamente os argumentos e os adjuntos que se apresentam numa estrutura sintagmática de tal maneira complexa que se caracterizam como orações: os argumentos e adjuntos oracionais. Trata-se, na verdade, de mais uma instância de nosso conhecido fenômeno da recursividade, a qual recebe bastante destaque nas aulas de língua portuguesa sob o nome de *orações subordinadas*. Este é o tópico da próxima seção.

A cisão dos sintagmas

Na literatura gerativista, desde o trabalho seminal de Pollock (1989), os sintagmas SC, ST e SV vem sendo reanalisados e descritos por meio de outros sintagmas mais específicos que os constituem. Assim, por exemplo, o SC pode ser dividido num "sintagma de tópico" (STop) e um "sintagma de foco" (SFoc). O ST pode ser reinterpretado como um sintagma de concordância para o sujeito (AgrS) e um sintagma de concordância para o objeto (AgrO) – note que *Agr* é o termo inglês que significa *concordância*. Também o SV vem sendo redescrito como a articulação entre um verbo leve ("V") – lido como *verzinho*, que dá origem ao Sv, que domina o SV. Na figura abaixo, você pode conferir a estrutura de uma sentença com o ST (TP) cindidos num sintagma de negação (NegP) e num sintagma de concordância (AgrP).



Exemplo de um sintagma cindido de acordo com Pollock (1989).

A motivação para a cisão de sintagmas depende do objetivo específico do sintaticista, em sua tarefa de explicar um determinado fenômeno. Pollock, por exemplo, teve a intenção de indicar que o sintagma de tempo comportava, na verdade, mais do que somente a expressão do tempo verbal. Ele dava conta também da negação, da concordância, do aspecto, dentre outras noções funcionais.

Argumentos e adjuntos oracionais

Conforme nos ensinam as melhores gramáticas tradicionais: “Quando a estrutura sintática de um período compreende uma única oração, dizemos que esse é um *período simples* e sua única oração é uma *oração absoluta*. Já quando ocorrem no período duas ou mais orações, dizemos que esse é um *período composto*”. Essas são definições bem abstratas e tradicionais, mas depois de tudo o que até aqui já estudamos, será fácil compreender o que querem dizer. Senão, vejamos.

Em (24) a seguir, temos uma frase do português, já que se trata de um enunciado linguístico comunicativo nessa língua natural.

(24) João fez uma bobagem na frente de todos.

Tal frase possui uma característica específica: ela possui uma estrutura verbal, identificada pela ocorrência do predicador “dizer”. Possuir um verbo predicador faz com que essa frase alcance o *status* de período. Isso quer dizer que o exemplo (24) não é apenas uma frase qualquer, mas sim uma *frase verbal*.

Mas o que acontece quando uma frase não possui predicado verbal? A resposta é a seguinte: na ausência de predicadores verbais, um enunciado linguístico é apenas uma *frase nominal*, tal como “Silêncio”, “Olá”, “Boa tarde”, “Socorro”, uma pequena oração ou qualquer outra expressão simples. Sobre esses tipos de frase, há muito pouco a se dizer no estudo da sintaxe. Logo, como já sabemos, o interesse do sintaticista recai especialmente sobre frases verbais, ou seja, sobre períodos.

Depois de entendermos o que é um período e conseguirmos identificá-lo numa frase, devemos seguir em frente com nossa análise linguística e descrever a sua estrutura interna. Devemos, portanto, identificar os predicadores verbais que, no interior do período, selecionam argumentos e recebem adjuntos. Voltando ao exemplo (24), percebemos que nele há somente um predicador verbal: o item “fazer”, que seleciona seu argumento externo (“João”), seu argumento interno (“uma bobagem”) e recebe um adjunto (“na frente de todos”). Já aprendemos que cada conjunto de relações entre um predicador verbal, seus argumentos e seus eventuais adjuntos é aquilo que tradicionalmente denominamos orações. Isso significa que, no período em (24), possuímos uma e somente uma oração.

Ora, quando isso acontece, isto é, quando encontramos períodos que são organizados em torno de uma e somente uma oração, dizemos tratar-se de um período simples, com uma oração absoluta (uma oração única na frase, tal como (24)). Em contrapartida, quando encontramos duas ou mais orações na estrutura de um período, dizemos se tratar de um período composto. Em (25), notamos a ocorrência de dois predicadores verbais, razão pela qual esse período deve ser identificado como composto, às vezes também denominado como *período complexo*.

(25) Paulo disse que João fez uma bobagem na frente de todos.

Nessa frase, verificamos a ocorrência do predicador verbal “fazer”, que seleciona dois argumentos e recebe um adjunto, conforme já descrevemos acerca de (24). Encontramos também o predicador verbal “dizer”, que seleciona “Paulo” como seu argumento externo e, como argumento interno, seleciona toda a oração “João fez uma bobagem na frente de todos”.

O interessante a respeito dos períodos compostos é que eles são a expressão máxima da recursividade das línguas naturais. Com eles, nos damos conta de que uma oração, depois de constituída por meio das complexas combinações de sintagmas que estudamos nesta unidade, poderá ser selecionada como constituinte de uma outra oração, seja como argumento ou como adjunto. Analise bem os exemplos (24) e (25) e você entenderá o que estamos dizendo. Note que a oração em (24) é um dos constituintes do período complexo apresentado em (25).

Ora, a diferença fundamental entre período simples e período composto é que, no período simples, os argumentos e os adjuntos associados a um dado predicador apresentam-se na forma de sintagmas (e não de orações), ao passo que, no período composto, argumentos e adjuntos apresentam-se na forma de orações (e não de sintagmas simples).

Talvez você se pergunte como diferenciamos as orações selecionadas como argumento das selecionadas como adjunto. Saiba que não há nada de novo ou especial nesse tipo de distinção. Argumentos e adjuntos distinguem-se sempre pelo fato de serem ou não selecionados pela estrutura argumental de um predicador, seja no período simples ou no composto. Dessa maneira, se uma oração ou um sintagma simples é selecionado por um predicador, então será seu argumento. Já se uma oração ou sintagma simples não é selecionado por um predicador, mas meramente a ele é adjungido, então se trata de um adjunto.

As orações que são selecionadas como argumento de algum predicador são tradicionalmente denominadas *orações substantivas*. Por exemplo, toda a oração “João fez uma bobagem na frente de todos” é selecionada como argumento interno do predicador “dizer” em (25). Sendo assim, essa é uma oração substantiva – isto é, uma *oração argumento*. Quando, numa nomenclatura por vezes excessiva e redundante, encontramos nos livros e nas aulas de português referências às orações subordinadas substantivas com suas diversas funções sintáticas, o que estão tentando nos dizer é que essas orações são o argumento de algum predicador.

Para facilitar a sua memória em relação às funções sintáticas do período, que tanto estudamos em escolas tradicionais, elaboramos a tabela a seguir. Note que a classificação oficial de uma dada função quando o respectivo constituinte se encontra em forma de oração é desnecessariamente grande. Por exemplo, poderíamos dizer que um "sujeito" em forma de oração é simplesmente um "sujeito oracional", em vez de dizermos "oração subordinada substantiva (O. S. S.) subjetiva".

Estrutura sintagmática	Função sintática	Quando oracional
Argumento externo	Sujeito	O. S. S. subjetiva
Argumento interno V,	Predicado	
quando SN	Objeto direto	O. S. S. objetiva direta
Argumento interno de V,	Objeto indireto	O. S. S. objetiva indireta
quando SP	Complemento Nominal	O. S. S. completiva nominal
Argumento interno de N ou A	Predicativo	O. S. S. predicativa
Argumento nominal	Agente da passiva	O. S. S. agentiva
Argumento em voz passiva		
Argumento a SV	Adjunto adverbial	O. S. Adverbial
Argumento a SN	Adjunto adnominal	O. S. Adjetiva

A função a seguir não aparece na nomenclatura escolar oficial, mas é bastante útil para diferenciarmos o argumento interno de verbos como [ir] do de verbos como [precisar].

Argumento interno de V (verbo de movimento espacial)	Complemento circunstancial
---	----------------------------

Se uma oração é apenas adjungida a algum sintagma, mas não é selecionada pela estrutura argumental de um predicador, então temos um adjunto oracional. Um adjunto em forma de oração — ou uma *oração adjunto* — pode ser classificado como oração adjetiva ou oração adverbial. A distinção decorre do tipo de sintagma que recebe a oração adjunto. Se o sintagma que recebe o adjunto é um SN, como vemos em (26), então dizemos que a oração é adjetiva, também chamada de *oração relativa* (note que, a seguir, usaremos dois colchetes contíguos para representar a ocorrência de uma adjunto sintático. Quando o constituinte representado é um complemento, usamos apenas um colchete simples).

(26) O [_{SN} livro [_{SC} que [_{ST} todo mundo leu]]] é muito bom.

Note que a oração [que todo mundo leu] é inserida na frase como um adjunto do nome [livro]. Modificadores nominais são chamados de adjetivos e, como esse adjetivo específico encontra-se em forma de oração, chamamo-lo de oração adjetiva ou adjunto adnominal oracional.

Por sua vez, as orações adverbiais são modificadores do verbo. Elas são inseridas na frase, portanto, como adjuntos de um SV, conforme vemos a seguir.

(27) Paulo [_{SV} chegou em casa [_{SC} quando [_{ST} começou a chover]]]

Você deve notar que a oração [quando começou a chover] insere-se na estrutura do período como modificador do SV [chegou em casa]. Como o próprio nome nos indica, "advérbios" são modificadores de verbos ou sintagmas verbais. No caso, o advérbio encontra-se na forma de uma oração inteira — e é por isso que o denominamos oração adverbial, ou adjunto adverbial oracional.

Como você pode ver, não há nada de excepcional no estudo do período composto. Ele é, como dissemos, apenas uma instância do fenômeno da recursividade, tão natural às línguas humanas. A razão para esse assunto receber tanta atenção em nossa formação escolar reside no fato de argumentos e adjuntos oracionais serem estruturas bastante complexas quando comparadas a argumentos e adjuntos não oracionais.

Uma vez que tenhamos consciência da recursividade das combinações sintáticas, já podemos esperar que orações possam ser combinadas com outras orações, tanto quanto sabemos que sintagmas podem se combinar com sintagmas, palavras podem se combinar com palavras, morfemas podem se combinar com morfemas, sílabas podem se combinar com sílabas, fonemas pode se combinar com fonemas... Essa é a natureza combinatoria recursiva que faz das línguas naturais um instrumento de comunicação e expressão tão poderoso e tão belo.

O tipo mais simples de recursividade existente numa língua natural é a mera justaposição de estruturas, isto é, a colocação lado a lado de elementos linguísticos. Por exemplo, se temos o nome "João", podemos combiná-lo com outro nome, coordenando-os por meio de uma conjunção como "e": [João e Paulo]. Dessa forma, as orações coordenadas que estudamos na escola devem ser interpretadas como uma expressão simples da recursividade das línguas. Quando coordenamos orações, simplesmente justapomos uma oração ao lado da outra, num dado período, usando ou não conectivos explícitos. É isso o que acontece em [[João saiu], [eu fiquei em casa]] e [[Paulo gosta de rock], [mas eu gosto mesmo de MPB]].

Conclusão

Nesta unidade, exploramos o arsenal analítico básico necessário ao estudo do funcionamento do Sistema Computacional da linguagem humana. Vimos que a noção fundamental da sintaxe é o sintagma. Este é entendido como o resultado da combinação binária entre dois constituintes. Se juntarmos a simplicidade da

computação binária que constrói um sintagma com a propriedade fundamental da recursividade, chegaremos à construção de estruturas tão intrincadas como a oração e o período composto. Dominar a noção de sintagma, diferenciar sintagmas lexicais e funcionais, identificar a fronteira entre orações etc. são apenas instrumentos metalinguísticos úteis e imprescindíveis para compreendermos como é que a nossa cognição produz e compreende frases no nosso uso cotidiano da linguagem. Esperemos que você tenha já o domínio dessas ferramentas. Você fará uso corrente delas na próxima unidade e durante toda a sua vida acadêmica como estudioso da sintaxe, caso deseje prosseguir com seus estudos.

Exercícios

- 1) O que se deve entender por "sintaxe" nos estudos de linguística gerativa?
- 2) Quais são as unidades mínimas e máximas da análise sintática?
- 3) O que se deve entender pela noção de sintagma?
- 4) Use os testes de identificação de sintagma para descrever a ambiguidade estrutural presente na seguinte frase.
"Recebi uma fotografia de Petrópolis."
- 5) O que os sintaticistas querem dizer quando afirmam que a constituição de sintagmas é sempre binária e pode ser recursiva? Dê um exemplo.
- 6) Identifique os sintagmas presentes nas estruturas que se seguem.
 - a) Lemos muitos textos.
 - b) Lemos muitos textos de linguística.
 - c) Lemos muitos textos de linguística gerativa.
- 7) Represente os seguintes sintagmas em árvores sintáticas.
 - a) Vi televisão.
 - b) João gosta de doces.
 - c) Paulo dormiu por horas.
- 8) Descreva os três sintagmas funcionais apresentados nesta unidade. Dê um exemplo de cada.
- 9) Como você compreendeu a diferença entre períodos simples e períodos compostos a partir desta unidade? O que essa diferença tem a ver com o fenômeno da recursividade?

UNIDADE 9

O Sistema Computacional

Na unidade anterior, aprendemos os fundamentos da análise sintática gerativista. Estudamos os diferentes tipos de sintagma existentes nas línguas naturais e aprendemos a desenhar suas relações estruturais por meio dos diagramas arbóreos — as árvores sintáticas. Todo esse instrumental técnico que acabamos de estudar possui uma função bem clara no entendimento da linguística gerativa: trata-se de ferramentas que nos ajudam a compreender e representar o conhecimento sintático que inconscientemente dominamos quando falamos uma dada língua. Isso quer dizer que precisamos saber identificar sintagmas, desenhar árvores e delimitar orações num dado período porque essa é uma habilidade muito útil, indispensável ao sintaticista que assume a tarefa de desvendar os mecanismos cognitivos por meio dos quais os humanos são capazes de produzir e compreender frases.

É correto dizer que tudo o que aprendemos na unidade 8 foi descrever formalmente as estruturas sintáticas da linguagem humana. Pois bem, na presente unidade, passaremos da descrição linguística *formal* para a descrição computacional das operações sintáticas subjacentes ao nosso uso da linguagem. Chegamos, portanto, ao momento do curso em que faremos uso do que já aprendemos sobre léxico e sintaxe para começar a analisar o Sistema Computacional da linguagem humana.

Dizemos que a sintaxe gerativa é uma abordagem *formal* acerca da linguagem humana porque se ocupa fundamentalmente das "formas" que estruturam fonemas, morfemas, palavras, e frases — independentemente das diversas *funções* que essas formas podem assumir no uso da língua. Os manuais de linguística normalmente operam à abordagem "formal" as abordagens "funcionais" que, em vez da "forma", privilegiam as "funções" comunicativas da linguagem.

Como já sabemos, o Sistema Computacional é somente um dentre os diversos componentes da linguagem humana. Léxico, Forma Fonética (FF), Forma Lógica (FL) e sistemas de interface (conceitual-intencional e articulatório-perceptual) são outros elementos no complexo edifício de nosso conhecimento linguístico. Cada

