

Morfologia e Sintaxe

Tradicionalmente morfologia e a sintaxe são respectivamente os domínios da palavra e da frase. A noção de palavra que, como veremos, escapa a qualquer definição formal unitária, está no centro dos estudos da linguagem desde a antiguidade. Veja-se essa definição do discurso (*oratio*), pelo gramático latino Diomedo¹: *‘Oratio est compositio dictionum consumans sententiam remque perfectam significans ... Oratio est ore missa et per dictiones ordinata pronuntiatio’* (O discurso é uma composição de palavras que completa um pensamento e significa uma realidade completa O discurso é produzido pela boca e a pronuncia é organizada em palavras). Para Diomedo, a palavra é uma *vox articulata cum aliqua significatione*, “um som articulado com alguma significação”.

As gramáticas escolares modernas retomam essa definição quase literalmente. Cunha (1972:54), por exemplo, afirma que *‘uma PALAVRA é constituída de elementos materiais (vogais, consoantes, semivogais, sílabas, acento tônico) a que se dá um sentido e que se presta a uma classificação.’* Como exemplo, o autor cita a palavra ‘boi’, designativo de *‘um quadrúpede ruminante que serve para os trabalhos de carga e para a alimentação’*, definido-a como *‘um substantivo comum, concreto, primitivo, simples, masculino, singular; monossílabo, tônico, formado da consoante /b/ seguida do ditongo decrescente /oi/’*. Em Abaurre, M. L. et al (2003:156) encontramos a seguinte definição de palavra: *‘Palavra é uma unidade lingüística de som e significado que entra na composição dos enunciados da língua’*.

Já a noção de frase, fazendo referência a uma seqüência de palavras sintaticamente organizadas, é uma noção moderna. Como vimos acima, o termo *Oratio* (*logos* em grego – tradicionalmente traduzido pelo termo *Discurso*) remete à expressão de um pensamento completo, noção impossível de definir de maneira precisa. É

¹ Apud Charpin (1980), p. 26.

interessante notar que as definições de Frase nos manuais escolares também fazem referência à completude do sentido. Cunha (1972:85) define frases como *'enunciações de sentido completo, as verdadeiras unidades da fala e das quais nos servimos quando expressamos nossos pensamentos e sentimentos'*. Para Abaurre, M. L. et al (2003:156), a frase é definida como *'um enunciado lingüístico que, independente de sua estrutura ou extensão, traduz um sentido completo em uma situação de comunicação. Outra característica da frase é que adotamos uma entoação específica que marca seu início e seu fim.'* Para definir frase, Cegalla (1998) também recorre ao sentido: *'frase é todo enunciado capaz de transmitir, a quem nos ouve ou lê, tudo o que pensamos, queremos ou sentimos. Pode revestir as mais variadas formas, desde a simples palavra até o período complexo, elaborado segundo os padrões sintáticos do idioma.'*

A lingüística moderna, de maneira mais ou menos explícita, retoma essa caracterização operacional. O que diferencia as teorias é a relação que se estabelece na descrição e explicação dos fenômenos entre a forma da frase e o seu sentido, em particular no que diz respeito à sua inserção numa dada situação de comunicação. Na abordagem funcionalista, essa relação é determinante, e a sintaxe é vista como o reflexo das funções comunicativas veiculadas pela frase. Desse ponto de vista, forma e uso são indissociáveis na explicação dos fenômenos sintáticos. Na abordagem da Teoria da Gramática Gerativa, ao contrário, a sintaxe é um componente autônomo, com princípios próprios que independem do uso.

Escolhemos organizar este capítulo não pela apresentação dos diversos modelos, mas em torno das grandes questões levantadas pelo estudo da linguagem. Na Seção I, veremos como a lingüística moderna formula as questões levantadas pela definição tradicional da frase. Na seção II, definiremos gramática à luz da dicotomia Língua-Interna/Língua-Externa, apresentada com base na comparação entre o Português europeu e o Português brasileiro. Os universais da linguagem e os parâmetros de variação entre as

línguas serão o assunto da Seção III. Na seção IV, levantaremos a questão da variação intra-linguística, e sua relação com a mudança. Enfim, na seção V, focalizaremos a palavra e o lugar da morfologia na arquitetura geral da gramática.

I. A frase no discurso

1. Forma e função

O debate da primazia da forma sobre a função ou da função sobre a forma é antigo, e não se limita à lingüística². Mas a tendência em antagonizar as abordagens que privilegiam uma ou outra tem sido forte na lingüística moderna. Na abordagem funcionalista, a sintaxe é vista como o reflexo das funções comunicativas veiculadas pela frase. Desse ponto de vista, forma e uso são indissociáveis na explicação dos fenômenos sintáticos. Na abordagem da Teoria da Gramática Gerativa desenvolvida por Chomsky e seus seguidores, ao contrário, a sintaxe é um componente autônomo, com princípios próprios que independem do uso. Mas como veremos a seguir, mesmo na abordagem gerativista, a questão da relação entre forma e função, entre gramática e uso, entre estrutura e interpretação semântica, se constitui numa questão central nas diversas formulações do modelo ao longo dos últimos 50 anos.

Encontramos em Halliday uma definição esclarecedora da abordagem funcionalista da linguagem:

Uma abordagem funcional da linguagem significa, antes de tudo, investigar como a linguagem é usada: tentar encontrar para que propósitos a linguagem nos serve, e como nós somos capazes de realizar estes propósitos, falando e ouvindo, lendo e escrevendo. Mas também significa mais que isto. Significa buscar explicar a natureza da linguagem em termos funcionais: vendo se a própria linguagem tem sido moldada pelo uso, e se sim, de quais modos – como a forma da linguagem tem sido determinada pela função a que ela serve. (Halliday, *Explorations in functions of language*, 1973, p.7; apud Givon 2001).

² Para uma discussão mais detalhada dessa questão, ver Givon (2001), Cap. 1.

O objetivo último das análises é, assim, para os linguistas funcionalistas, encontrar na forma dos enunciados o reflexo da função comunicativa/discursiva que eles veiculam. E esse reflexo certamente existe. As línguas têm todas à sua disposição mecanismos para expressar o mesmo conteúdo semântico-proposicional (ou seja, a expressão de eventos e o estabelecimento de relações entre participantes nesse evento) – também chamado de estrutura lógica -, de diversas maneiras conforme diferentes organizações discursivas:

- 1) O gato comeu o rato.
- 2) O rato foi comido pelo gato.
- 3) Foi o gato que comeu o rato (e não o cachorro).
- 4) Foi o rato que o gato comeu (e não a lagartixa).

Todas as orações acima têm o mesmo conteúdo semântico proposicional, ou seja, dizem que houve um evento representável da seguinte maneira: “x comeu y” sendo que “x= o gato” e “y= o rato”. Mas elas se diferenciam pela maneira como o dizem. Assim a frase 1 diz alguma coisa do gato, enquanto que a frase 2 fala do rato. Em termos mais técnicos, diremos que o tema de 1) é o gato, e o tema de 2) é o rato. Isso aponta para contextos discursivos distintos. Já em 3) e 4), não há propriamente temas, já que não há asserções a respeito do rato ou do gato, mas a explicitação de que o agente do ato de comer foi o gato e de que o objeto comido foi o rato. O gato e o rato não são mais os temas, mas os focos - a informação nova – da asserção. Em 3) e 4), contrariamente a 1) e 2), a existência de um evento no qual o rato foi comido, ou o gato comeu alguma coisa, já é pressuposta. Essa informação já é compartilhada pelos falantes, nesse sentido é

informação velha. Informação velha e nova são noções que remetem ao contexto discursivo no qual os enunciados são produzidos, e vemos claramente que a sintaxe do português codifica essas noções. De fato todas as línguas têm recursos para codificar essas noções.³

A questão que divide as teorias é assim a centralidade ou marginalidade desse efeito do discurso sobre a forma dos enunciados no modelo de sintaxe que se constrói. Nas palavras de Givón, as diversas construções sintáticas não são senão a codificação das funções pragmático-discursivas:

Construções ou tipos de oração são então instrumentos de codificação gramatical que sinalizam funções discursivo-pragmáticas. (op. cit. p. 19).

Do ponto de vista da Teoria da Gramática Gerativa, essa idéia da gramática codificando as funções discursivo-pragmáticas é inaceitável porque o componente gramatical é autônomo em relação ao componente pragmático que, para Chomsky, faz parte dos sistemas de desempenho. Na última versão da teoria, o Programa Minimalista, a gramática é representada como um sistema gerativo que constrói objetos lingüísticos interpretados como instruções para os sistemas de desempenho:

A língua está encaixada em sistemas de performance que permitem que as suas expressões sejam usadas para articular, interpretar, referir, perguntar, refletir, e exercer

³ Note-se que não são forçosamente recursos sintáticos. A marcação fonológica existe em muitas línguas, inclusive no português, onde uma frase como (4) pode ser parafraseada por i):

i) ‘O gato comeu o **RATO** (e não a lagartixa).’
Onde as maiúsculas em i) representam um acento de intensidade na palavra “rato”.
Em certas línguas, o foco não é marcado apenas sintaticamente, mas também morfológicamente e fonologicamente. O exemplo ii) é um exemplo de marcação morfológica, sintática e fonológica em hindi-urdu (cf Kidwai, 1999). Em i), o elemento ‘Ram’ é focalizado sintaticamente, através da ocupação da posição pré-verbal, morfológicamente, pela focalização *in situ* via cliticização do elemento *hi* a este elemento, e fonologicamente, através do acento contrastivo que ‘Ram’ porta. Nenhuma destas estratégias de focalização se encontra em distribuição complementar em hindi-urdu, uma vez que todas as três podem ser utilizadas em uma única sentença.

ii) kitaab **Ram**-hii laayegaa (siitaa nahii).
livro **Ram**-Emph trará Sita não
‘**Ram** trará o livro, não Sita.’

outras ações. Podemos considerar que cada DE⁴ é um complexo de instruções para estes sistemas de performance, fornecendo informação relevante para o seu funcionamento. (Chomsky (1995), p. 168, tradução de Eduardo Raposo, p. 244).

Desse ponto de vista, a relação sintaxe/contexto discursivo é inversa àquela encontrada na concepção funcionalista da linguagem. No modelo chomskiano, a sintaxe vem primeiro, obedece a princípios próprios, e gera estruturas que funcionam como instruções para os sistemas de desempenho.

Contudo, Chomsky não nega, sobretudo nos seus últimos textos, uma relação mais orgânica entre forma e função da linguagem:

Se bem que a idéia de que a linguagem é “desenhada com vista ao uso” ou “bem adaptada às suas funções” não tenha um sentido claro, esperamos encontrar conexões entre as propriedades da linguagem e a maneira como é usada. (Op.cit., p. tradução de Eduardo Raposo, p. 244)

O lugar exato dessa conexão é que continua em debate na teoria.

Por outro lado, autores como Givón não negam que a linguagem comporta uma grande parte de arbitrariedade, ou seja, que nem tudo na língua pode ser explicado pela função desempenhada no discurso – o que ele chama de aspecto icônico da gramática:

“A iconicidade da linguagem não é absoluta, mas antes uma questão de grau. Na maior parte das construções gramaticais, recursos (ou princípios) mais icônicos são mesclados com recursos (ou regras) simbólicos convencionais mais arbitrários. (op. cit. p. 34)

Nas próximas seções, mostraremos, por um lado, como o estudo de diversos fenômenos dá uma justificação empírica à distinção entre níveis diferentes, articulados entre si, porém, de natureza diferente. Definiremos uma gramática de frase⁵, regida por

⁴ “DE” significa “Descrição estrutural”

⁵ Cabe aqui uma observação sobre a tradução do termo inglês “phrase”. O equivalente em português não é propriamente “frase”, mas “sintagma”. Nesse sentido, seria mais adequado falar em “gramática sintagmática”. Porém, uma vez que esse termo não é de uso corrente, continuaremos a usar o termo

princípios próprios, em interface com uma gramática de discurso, regida por outros princípios. Mostraremos como a primeira é logicamente anterior à segunda, uma vez que ela fecha opções. Mostraremos também que as línguas diferem entre si em relação àquilo que faz parte da gramática de frase, apontando para a necessidade de distinguir claramente dois níveis para dar conta de maneira satisfatória da diferença entre as línguas.

2. Dois níveis de interpretação semântica

A distinção entre a gramática de frase e gramática de discurso na gramática gerativa tem sua origem no modelo de interpretação semântica proposto por Chomsky no seu livro *Reflexions on Language*. Nesse livro, Chomsky retoma sua polêmica com os proponentes da Semântica Gerativa, que propunham que toda a interpretação semântica das frases fosse codificada na estrutura profunda a elas associada⁶.

Um bom exemplo dessa teoria se encontra na questão da interpretação dos pronomes. Na Semântica gerativa, um enunciado ambíguo como:

5) João sabe que ele é inteligente.

tem duas derivações diferentes conforme a interpretação do pronome. Na interpretação em que *ele* refere a uma pessoa previamente mencionada no discurso, distinta de *João*, a estrutura profunda já contém o pronome, ficando igual à estrutura superficial. Mas na interpretação em que *ele* refere a *João*, a estrutura profunda contém duas ocorrências de *João*: *João sabe que João é inteligente*. Uma regra de

‘gramática de frase’.

⁶ Os detalhes do debate se encontram nos ensaios reunidos em *Studies on Semantics in Generative Grammar*.

pronominalização transforma essa estrutura profunda na estrutura superficial 5) transformando a segunda ocorrência de *João* no pronome *ele*.

Para Chomsky, o que é relevante do ponto de vista da interpretação da frase é que, em 5), *ele* pode ser interpretado como *João*, mas não necessariamente⁷. Isso, crucialmente, faz parte da competência de qualquer locutor nativo de português, e contrasta a interpretação de 5) com a de 6):

6) Ele sabe que João é inteligente

Em 6), *ele* não pode ser *João*. Chomsky distingue então o fato gramatical do fato discursivo. A gramática nos diz que em 5) *ele* pode ser *João*, e que em 6), não pode. Saber se *ele* é ou não é *João* em 5) depende de um conhecimento outro, que remete ao contexto discursivo no qual a frase foi enunciada.

A partir dessa distinção, Chomsky propõe um modelo de gramática no qual a semântica não é gerativa, mas interpretativa, ou seja, a sintaxe não é gerada a partir da semântica, mas a semântica interpreta a sintaxe. Isso dá origem ao modelo em T onde a estrutura sintática gerada pelas regras transformacionais constitui a interface entre as regras de interpretação semântica e as regras de interpretação fonológica.⁸

Além disso, a interpretação semântica como um todo se divide em dois componentes distintos. Distinguem-se as regras de interpretação semântica (IS) que pertencem à gramática de frase (IS-1), e as que pertencem a outros sistemas em articulação com a gramática de frase (IS-2). O que distingue essas regras não é só que as regras do componente IS2 dependem efetivamente do contexto discursivo, mas também

⁷ Um outro argumento contra esse tipo de análise é de natureza meta-teórico: uma derivação como a proposta acima torna o modelo muito custoso, uma vez que as regras envolvidas são muito poderosas: elas podem modificar drasticamente o material lexical envolvido, transformando o nome próprio em pronome.

⁸ De fato o modelo em T só parece explicitamente no livro fundador da teoria de Princípios e Parâmetros: *Lectures on Government and Binding*. Este modelo será apresentado na seção I.4:

que as regras do componente IS1 são *dependentes da estrutura*, como os outros fenômenos sintáticos⁹. Veja-se de novo o exemplo acima. Uma análise mais detalhada mostra que o que está em jogo nas possibilidades de identidade referencial entre o nome e o pronome não é a sua posição relativa na ordem linear do enunciado (num caso o nome precede o pronome, no outro o segue), mas uma relação estrutural mais complexa. Com efeito, numa frase como 7):

- 7) A insinuação de que ele poderia perder o emprego deixou o João muito abalado.

ele e *João* podem ser a mesma pessoa apesar de *ele* preceder *João*. O fato é que há uma diferença crucial entre (6) e (7). Em (6), o pronome é o sujeito da frase principal, em (7), ele é o sujeito de uma frase encaixada dentro do sujeito da frase principal, como representado a seguir¹⁰:

- 6') [**ele** [sabe que João é inteligente]]
- 7') [[a insinuação [de [que [**ele** poderia perder o emprego]]]] [deixou o João muito abalado]]

Em termos técnicos, o pronome c-comanda¹¹ o nome em 6), mas não em 7). O c-comando é um princípio que restringe as relações sintáticas fundamentais, definindo as

⁹ A interpretação dos pronomes, como de todas as expressões nominais, é regida pelos *Princípios de Ligação*. A ligação inclui uma condição estrutural, chamada *Condição de c-comando*, ativa sempre que uma relação de natureza sintática existe entre dois elementos da frase, como por exemplo a atribuição de caso, ou a relação entre um elemento deslocado e sua posição de origem..

¹⁰ Usamos aqui colchetes para representar a estrutura sintagmática da frase. Essa representação pode ser traduzida numa representação arbórea, onde cada par de colchetes corresponde a um nó da árvore.

¹¹ A c-comanda B se, e somente se, A não domina B e a primeira categoria ramificante que domina A domina B. As relações de dominância estão representadas na representação em colchete pela contiguidade com a fronteira marcada em negrito. Note-se que por essa definição, o sujeito de uma sentença, que é imediatamente dominado pela categoria correspondendo a essa sentença, c-comanda todo o resto da sentença. Já o sujeito em 7) é imediatamente dominado por uma categoria (em negrito na representação) que não domina o nome *João*.

configurações possíveis nas quais dois elementos da estrutura entram numa determinada relação. Ele também desempenha um papel central na interpretação dos pronomes, e das expressões nominais em geral, entrando na definição de *ligação* (em inglês “binding”). Uma determinada expressão nominal A liga outra expressão nominal B quando, além de haver co-referência entre A e B, A c-comanda B. Nesse caso, se configura uma dependência referencial de B em relação a A. No caso de 6), uma vez que *ele* c-comanda *João*, se *ele* e *João* são co-referentes, *ele* “liga” João, ou seja, *João* depende referencialmente de *ele*. Isso viola um dos princípios de ligação que diz que os nomes – tendo uma referência intrínseca - não podem ser ligados¹². Em 7), *ele* não c-comanda *João*, e, portanto, não está em situação de ligar *João*. É o que permite, de maneira um tanto paradoxal, que *ele* e *João* sejam co-referentes. De fato, a correferência por si só não é uma noção gramatical, mas mais bem uma noção discursiva. Saber se *ele* e *João* são a mesma pessoa em 7) não depende da gramática. O que a gramática diz é que podem ser a mesma pessoa. Já em 6), isso não é possível porque, se fossem co-referentes, a relação estrutural em que se encontram definiria uma relação de dependência referencial que viola os princípios de ligação. Esse bloqueio da interpretação de identidade referencial entre o pronome e o nome aponta para a precedência da gramática sobre o discurso.

3. Gramática de frase e gramática de discurso

Vimos acima que a frase é de fato o âmbito de dois tipos de regras de interpretação semântica. As que são regidas por princípios internos, sensíveis à estrutura, e outras que são de natureza diferente, e remetem ao contexto discursivo.

¹² Os princípios de ligação são três:

- O princípio A, que diz respeito às anáforas (pronomes reflexivos e recíprocos), diz que uma anáfora tem que ser ligada na oração que a contem imediatamente;
- O princípio B, que diz respeito aos pronomes pessoais, diz que um pronome pessoal não pode ser ligado na oração que o contem imediatamente;
- O Princípio C, que diz respeito aos nomes, diz que um nome não pode ser ligado nunca.

Vejam os outros exemplos dessa distinção, que dizem respeito, de novo, à interpretação dos pronomes. Os princípios de ligação bloqueiam também a interpretação de *ele* como *Pedro* na frase seguinte:

- 8) *O Pedro_i viu ele_i

Já em 9), essa interpretação é lícita, apesar de não obrigatória:

- 9) O Pedro sabe que a Maria viu ele.

Isso se deve ao fato de que o princípio de ligação que diz respeito aos pronomes é sensível a um outro aspecto típico da gramática: a localidade. Um pronome não pode ser ligado por outra expressão nominal na oração que o contém, mas o pode fora dela. A diferença entre 8) e 9) é que em 8) *o Pedro* e *ele* estão na mesma oração, mas não em 9).

Consideremos agora a interpretação do pronome *ele* nas frases 10) e 11):

- 10) O Zé, o Pedro viu ele.
11) O Zé, o Pedro sabe que a Maria viu ele.

Tanto em 10) quanto em 11), a única interpretação possível para *ele* é *Zé*. Pelo que vimos até agora, isso não decorre imediatamente dos princípios de ligação. Com efeito, esses têm como efeito impedir ou tornar possível uma relação de correferência. Mas nunca a tornam obrigatória. Por outro lado, pelos princípios de ligação, *ele* não poderia ser interpretado como referindo a *Pedro* em 10), mas poderia em 11). Ora essa interpretação é bloqueada nos dois casos. O que está em jogo aqui é claramente um outro princípio. O sintagma nominal *O Zé* é, nos dois casos, o tópico discursivo do qual as orações *o Pedro viu ele* e *o Pedro sabe que a Maria viu ele* são o comentário. Isso

implica que haja nessas orações uma posição ocupada por uma expressão nominal referindo a *Zé*. O pronome desempenha essa função. Esse é um princípio de natureza discursiva que pode ser derivado de um princípio maior de coerência discursiva.

A comparação entre 8) e 9) por um lado, e 10) e 11) pelo outro, faz aparecer uma outra diferença entre os fenômenos de frase e os fenômenos de discurso. No primeiro caso, existe a sensibilidade já apontada para a localidade da relação. Faz diferença o pronome estar ou não estar na mesma oração que o nome. Já em 10) e 11), isso claramente não interfere. É porque a relação do tópico com o pronome não é direta, mas mediatizada pela sentença interpretada como comentário. Essa relação está fora do escopo dos princípios de ligação. Ela é de mesma natureza que a que se estabelece entre duas frases sucessivas:

- 12) O *Zé* não pode vir amanhã. Ele tem um compromisso importante em São Paulo.

Em 12), é um princípio de coerência discursiva que nos leva a interpretar o pronome sujeito da segunda sentença como sendo *Zé*. Obviamente, nesse caso, a gramática de frase não tem nenhuma incidência, uma vez que o nome e o pronome não estão na mesma frase. Isso nos leva a pensar que o tópico em 10) e 11) também está de alguma maneira externo à frase. Isso explicaria um fenômeno ilustrado pelo seguinte contraste do português europeu apresentado por Inês Duarte na sua tese de doutorado sobre o tópico em português europeu:

- 13) O João, imagina que o amigo dividiu com ele os direitos de autor
14) *Imagina que o João, o amigo dividiu com ele os direitos de autor!

A diferença entre 13) e 14) reside no fato de o tópico *O João* estar no início da sentença em 13), ao passo que em 14) ele está seguindo a conjunção *que* e precedendo imediatamente a oração subordinada. Segundo Duarte, 13) é uma frase bem formada, mas 14) não. Segundo Cinque (1983), esse contraste é encontrado em outras línguas românicas como o italiano, e

“... a regra responsável pela conexão entre o sintagma nominal e o pronome independente não é parte da gramática de frase, mas de um princípio da gramática discursiva, o mesmo que intervem entre sintagma nominal pleno e um pronome em duas frases adjacentes” op. cit. p. 13.

Ou seja, 14) é bloqueado pela impossibilidade de inserir uma relação de tipo discursivo no interior de uma oração. Isso, contudo, não significa que um tópico não possa ser encaixado numa oração. Observe-se as seguintes frases, ambas dadas por Duarte como bem formadas em português europeu:

15) Ao teu amigo sabes se já lhe pagaram os direitos de autor?

16) Sabes se ao teu amigo já lhe pagaram os direitos de autor?

As diferenças entre 13 e 14 de um lado, e 15 e 16 do outro, são duas. Nestas, por um lado, o elemento que retoma o tópico não é o pronome pleno *ele*, mas o clítico *lhe*. Por outro lado, existe o que Duarte chama de “conectividade casual” entre o tópico e o pronome: os dois têm marca de dativo; o clítico, morfologicamente e o tópico, pela presença da preposição *a*. Nesse caso, não há mais restrição sobre a posição do tópico. Ele pode tanto iniciar a frase toda, quanto preceder imediatamente a oração subordinada. Isso mostra que a relação entre o clítico e o sintagma nominal tópico é de outra natureza, regida pela gramática de frase. Isso é confirmado pelo fato, observado por Duarte, de que essa relação é submetida a restrições de localidade, como ilustrado em 17):

- 17) *Ao teu amigo, conheço um editor que ainda não lhe pagou os direitos de autor.

Em 17), o pronome clítico *lhe* está inserido numa oração relativa. Essas orações costumam constituir o que o linguísta Haj Ross chamou de ilhas, ou seja, elas não permitem a extração de algum constituinte para fora dos seus limites. A relação entre o pronome e o sintagma nominal em 17) sofre uma restrição similar. Não se pode estabelecer a relação de identidade referencial por cima dos limites da relativa. Isso é típico das relações definidas pela gramática de frase.

O leitor brasileiro já terá reparado que o paradigma do português europeu (doravante PE) descrito acima não se verifica no português brasileiro (doravante PB). Em PB, não há nenhum problema com frases como 14). Vários estudos mostraram que elas são perfeitas, como ilustrado pelos exemplos abaixo respectivamente tirados de Duarte (1995) e Kato (1993).

- 18) Eu acho que o povo brasileiro, ele tem uma grave doença.
19) Pedro pensa que essas crianças, a Maria esqueceu de pegar elas na escola.

Kato (1993) também mostra que o tópico pode ser encaixado numa oração relativa:

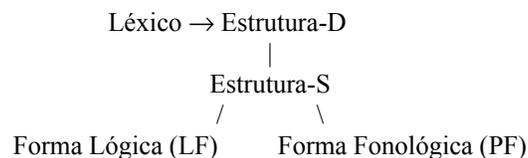
- 20) Esse país que o presidente, o povo não acredita mais nele, parece que saiu do marasmo.

Isso mostra que as línguas diferem em relação àquilo que faz parte da gramática de frase. O PB é uma língua na qual o tópico parece integrado à gramática de frase. A

primeira a apontar para essa característica foi Eunice Pontes, num trabalho que deu início a uma linha de pesquisa muito frutífera.¹³ Com base nos estudos tipológicos de Li e Thompson (1976), Pontes mostrou que o PB tinha um funcionamento sintático de “língua orientada para o tópico”, aproximando-se nisso de línguas geneticamente muito distintas como o chinês. Essas línguas têm como particularidade privilegiar a relação tópico/comentário na frase. Voltaremos a essa característica do PB na seção II, ao discutir a noção de “Língua-Interna”. Antes disso, faremos uma breve apresentação da arquitetura da gramática, ou seja, da maneira como os seus diversos componentes se relacionam no modelo de *Princípios e Parâmetros*¹⁴ da Teoria da Gramática Gerativa.

4. A arquitetura da gramática

No modelo de Princípios e Parâmetros a gramática pode ser representada da seguinte maneira:



O léxico é o ponto de partida de toda a derivação. Nele, os itens lexicais estão listados não só com os seus traços semânticos e fonológicos, mas também com as suas propriedades de “seleção”. A seleção pode ser unicamente sintática, quando envolve categorias que não têm propriedades semânticas, mas pode ser também de natureza

¹³ Cf. Galves (1998) e os ensaios sobre o PB, reunidos em Galves (2001), e Negrão (1999), entre outros.

¹⁴ O lugar da morfologia será mais sistematicamente tratado na seção V. Para uma introdução em português ao Modelo de Princípios e Parâmetros da Teoria da Gramática Gerativa, ver Raposo (1992), e Miotto et al. (2004). O livro de Lúcia Lobato é mais antigo, mas apresenta uma interessante retrospectiva da Teoria Gerativa, além de uma história mais geral da reflexão grammatical.

semântica. Nesse caso, se fala em “estrutura argumental”. A noção de estrutura argumental tem uma relevância particular para os verbos, que é a categoria com mais diversidade na definição de elementos selecionados¹⁵. A tradicional questão de classes de verbos pode ser traduzida em termos de estrutura argumental. Um verbo transitivo, por exemplo, tem dois argumentos, seu complemento, ou argumento interno, e seu sujeito, ou argumento externo¹⁶. Os verbos bitransitivos têm três argumentos, dos quais dois são argumentos internos, os intransitivos só têm argumento externo. Faz parte da competência lexical dos falantes associarem a cada verbo sua estrutura de argumentos.

A estrutura-D (do inglês *Deep Structure*, ou Estrutura profunda) é a projeção direta da estrutura argumental. Nesse nível, os argumentos estão na posição definida pelo léxico. Mas essa posição não é forçosamente aquela que aparece visivelmente na estrutura. Veremos vários casos dessa não correspondência da posição sintática dos elementos da frase e da sua interpretação mais abaixo. Nas orações passivas, por exemplo, o que é interpretado como o complemento do verbo está na posição sintática de sujeito. As orações interrogativas também costumam ter argumentos deslocados no início da frase, independentemente da posição onde estão interpretados. O deslocamento é uma propriedade das línguas humanas. Por isso, existe um nível adicional de representação, relacionado com a Estrutura-D pela operação de movimento: a Estrutura-S (ou Estrutura Superficial)¹⁷. A Estrutura-D e a Estrutura-S estão relacionadas uma a outra pelas operações de movimento¹⁸.

¹⁵ A noção de argumento é emprestada da lógica clássica, onde predicados são saturados por argumentos.

¹⁶ Para argumentos de que os complementos são mais internos do que os sujeitos, cf. Miotto et al., Cap. III.

¹⁷ A superficialidade da Estrutura-S é obviamente muito mais relativa do que no Modelo Padrão da Gramática Gerativa. A rigor, a estrutura superficial nesse modelo é a forma fonológica.

¹⁸ Note-se que não são só os argumentos que se movem, mas também outros elementos, como os verbos que se deslocam do sintagma verbal para se juntarem aos seus elementos de flexão. Isso é parametrizado nas línguas. Há evidências de deslocamento do verbo em línguas onde certos elementos, como a negação e certos advérbios, sempre seguem o verbo, por oposição a outras, onde esses elementos sempre precedem o

Nesse modelo, a estrutura-S é o nível de representação sobre o qual incidem tanto as regras de interpretação fonológica, quanto as regras de interpretação semântica. Os dois níveis Forma Lógica¹⁹ (ou LF, do inglês *Logical Form*) e Forma fonológica (ou PF, do inglês *Phonological Form*) são a representação, respectivamente, do som e do sentido da frase, tal como são definidos pela gramática. São esses dois níveis que estão em interface com os módulos de desempenho. No que diz respeito à interpretação semântica, as regras IS1 de *Reflexion on Language* são as regras que associam a Estrutura-S à Forma Lógica. As regras IS-2 são todas aquelas que remetem à coerência discursiva, ao conhecimento do mundo, ao sistema de crenças, que compõem o que Chomsky, mais recentemente, chamou de Módulo Conceptual-Intencional. Paralelamente ao módulo Conceptual-Intencional que interpreta a Forma Lógica gerada pela gramática, temos o módulo articulatorio perceptual que interpreta a Forma Fonológica. Nos dois casos, para dar conta dos fenômenos de linguagem, o modelo chomskiano define um espaço interno à gramática, e um espaço externo a ela, porém, em articulação com ela²⁰.

II. Língua interna/Língua externa

O que é uma língua? Em *Knowledge of Language*, Chomsky distingue duas definições: língua como Língua-Interna (Língua-I) e língua como Língua-Externa

verbo. Cf. o contraste entre o francês e o inglês representados respectivamente em (i) e (ii):

- (i) Jean voit souvent Marie.
- (ii) John *often* sees Mary.

¹⁹ O termo Forma Lógica se deve ao uso de recursos de notação da lógica clássica para a representação de certas relações semânticas.

²⁰ Cf. a citação de Chomsky (1995) na Seção I.1.

(Língua-E). A Língua-E pode ser definida como “a totalidade dos enunciados que podem ser produzidos numa comunidade de fala”. A Língua-I, ao contrário, é um objeto mental, o saber que as pessoas têm da língua e que lhes permite não só falar e entender essa língua, mas também:

“atribuir um estatuto a cada evento físico relevante, digamos cada onda sonora. Algumas são sentenças dotadas de sentido, literal ou figurado ... algumas são inteligíveis, tendo eventualmente um sentido definido, mas são mal-formadas de alguma maneira... algumas são bem-formadas, mas ininteligíveis. Algumas recebem uma interpretação fonética e mais nada; são identificadas como possíveis sentenças de uma língua, mas não a minha. Algumas são puro barulho. Há muitas possibilidades. Cada língua-I atribuirá um estatuto a cada um desses casos e a outros de maneira diferente.” (Chomsky 1985, p.26)

Desse ponto de vista, a gramática de uma língua é um objeto mental: o saber lingüístico dos falantes que têm essa língua como língua materna, desenvolvido durante o processo natural de aquisição. É o que faz com que um falante (inconscientemente) atribua estruturas aos enunciados que ele produz e ouve. A descrição dessas estruturas e, portanto, a caracterização dessa gramática é o objeto de investigação do lingüista trabalhando nesse quadro teórico.

A comparação do português europeu e brasileiro traz uma boa ilustração da relevância dessa abordagem para a compreensão dos fatos sintáticos das duas línguas numa abordagem comparatista. Desde o séc. 19 surgiu a questão da identidade do português falado no Brasil e em Portugal. Defensores das duas posições “é a mesma língua”/“não é a mesma língua” têm argumentos empíricos de peso ao apontar respectivamente para as semelhanças e as diferenças entre as duas vertentes. Como medir objetivamente o que faz pender a balança num sentido ou no outro? A questão se coloca de maneira diferente se em lugar de olharmos simplesmente para os enunciados, colocarmos a questão da estrutura associada a esses enunciados pelos falantes, ou seja, se colocarmos a questão não mais em termos de Língua-E, mas em termos de Língua-I.

Com efeito, desse ponto de vista, é possível que *dois enunciados superficialmente iguais tenham estruturas subjacentes diferentes* conforme são produzidas/interpretadas por um falante português ou brasileiro. Obviamente, damos então um passo na abstração, própria da abordagem mentalista, que deverá assentar numa argumentação empírica. Consideremos o seguinte enunciado:

21) O João é difícil de pagar.

Para um falante de português brasileiro, a primeira frase é ambígua: ela pode receber as duas interpretações seguintes:

21a) É difícil pagar João (ele não aceita dinheiro facilmente).

21b) João paga dificilmente (ele não gosta de pagar).

As interpretações a) e b) mostram que existe a possibilidade de interpretar *João* seja como o objeto de *pagar*, seja como o sujeito de *pagar*. No primeiro caso, o sujeito de pagar é arbitrário (qualquer pessoa). No segundo caso, é o objeto que é arbitrário. Já para um falante de português europeu, só a primeira interpretação é possível. Ou seja, *João* só pode ser interpretado como o objeto de pagar e nunca como seu sujeito. Isso se explica se postularmos que as estruturas subjacentes atribuídas pelos falantes de PB e de PE à mesma frase são, pelo menos em parte, diferentes. Podemos representar essas estruturas, correspondendo às interpretações mencionadas, da seguinte maneira:

21a') João_i é difícil de cv pagar cv_i

21b') João_i é difícil de cv_i pagar cv

Na estrutura, temos elementos que não são visíveis no enunciado, porque não têm realização morfológica, mas que são interpretados pelos falantes. Representamos esses elementos por cv, *categoria vazia*. Aqui, podemos dizer que se trata de um nome nulo. Além disso, usamos o recurso da co-indexação para marcar a identidade referencial entre as posições: a cv interpretada como *João* em a' é o objeto, mas em b' é o sujeito. Se ficarmos na superficialidade do enunciado, não vemos nenhuma diferença entre o enunciado brasileiro e o enunciado português. A seqüência morfossintática é a mesma. Se passarmos para a estrutura, que representa a interpretação, e, por isso, contem elementos abstratos, percebemos que os falantes de PB e de PE não tratam o enunciado da mesma maneira. Crucialmente, os primeiros têm a sua disposição uma estrutura a mais para associar ao enunciado. Isso mostra que o “sistema subjacente de regras dominado pelo locutor-ouvinte” brasileiro, para retomar os termos de Chomsky, é diferente do “sistema subjacente de regras dominado pelo locutor-ouvinte” português. Ou seja, que suas gramáticas – ou Línguas-I – são diferentes. São muitos os outros casos na língua que nos levam à mesma conclusão. Apresentarei mais um, que ilustra o mesmo ponto de maneira um pouco diferente. Considere-se a seguinte frase:

22) O relógio quebrou o ponteiro.

Em PB, essa frase é sinônima de:

23) Quebrou o ponteiro do relógio.

Para os locutores de PE, essa interpretação é impossível. A frase fica, portanto, inaceitável, a não ser no contexto de um mundo de fantasia no qual o relógio tem a possibilidade de quebrar seu próprio ponteiro. Isso significa que a única interpretação disponível para os falantes portugueses é aquela em que o sujeito de *quebrar* é interpretado como agente da quebra, como em:

24) O João quebrou o vaso de flores.

Note-se que essa interpretação também é possível para os falantes brasileiros, mas ela é normalmente obliterada por razões discursivas: a não ser nos filmes de Walt Disney, um relógio não pode ser agente. Se distinguirmos bem esses dois planos, o discursivo e o sintático, percebemos que, de novo, a Língua-I dos locutores brasileiros possibilita duas interpretações onde a Língua-I dos locutores portugueses só disponibiliza uma. Temos, portanto, razões de pensar que são diferentes.

Para representar a interpretação de 22) pelos locutores brasileiros, temos de recorrer de novo a categorias vazias co-indexadas, que mostram as relações percebidas pelos falantes, apesar da ausência de elementos morfológicos visíveis expressando essa relação. A estrutura (22') representa a interpretação de *relógio* em 22) como complemento de *ponteiro*:

22') O relógio_i quebrou [o ponteiro cv_i]

A impossibilidade para um locutor de PE associar essa mesma interpretação ao enunciado mostra que essa estrutura não é disponível na sua língua-I. Podemos dizer que essa estrutura não é gramatical em PE. Vale ressaltar então que o que é ou não é gramatical não é um enunciado, mas uma estrutura abstrata associada a um enunciado.

Quanto ao enunciado, será aceitável ou não. Vimos acima que a sua aceitabilidade pode depender de considerações discursivas (por exemplo, da determinação de mundos possíveis). As noções de “aceitabilidade” e “agramaticalidade” devem, portanto, ser distinguidas. “Aceitabilidade” é uma noção que remete à Língua-Externa e ao uso. “Gramaticalidade” remete à estrutura, portanto à Língua-Interna e à competência.

Os fenômenos que apresentamos até agora não são isolados, o que é de se esperar se eles são produzidos por diferenças gramaticais. É uma característica do PB, por oposição ao PE, ter em posição inicial da oração um sintagma nominal que é interpretado em relação com uma outra posição:

- 25) Essa bolsa cabe muitas coisas.
 “Cabe muitas coisas nessa bolsa.”
- 26) O João operou.
 “Alguém operou o João.”
- 27) A revista xerocou
 “Alguém xerocou a revista.”

Note-se também que as frases 26) e 27) poderiam ser parafraseadas por uma passiva²¹:

- 28) O João foi operado.
- 29) A revista foi xerocada.

²¹ Note-se, contudo, que o agente está mais presente na interpretação da passiva. Assim, pode-se dizer (i) mas não (ii):

- (i) A revista foi xerocada para ganhar tempo.
- (ii) *A revista xerocou para ganhar tempo.

A diferença entre frases como 26-27 por um lado, e as passivas em 28) e 29), é que as primeiras não assinalam por nenhuma marca morfológica a modificação operada na expressão das relações semânticas na frase. Com efeito, como explicitado nas glosas de 26) e 27), os sintagmas nominais sujeito nessas frases são interpretados como sendo o complemento do verbo (seu argumento interno). É também o caso nas passivas. Mas nestas, a morfologia no verbo e o uso do auxiliar *ser* assinalam essa alteração na projeção sintática dos argumentos do verbo. É uma característica das chamadas línguas de tópico poderem colocar muito livremente um sintagma complemento em posição inicial de frase, onde ele adquire estatuto de sujeito, sem que haja nenhuma marca morfológica no verbo marcando a mudança de função gramatical do argumento. Isso é obviamente impossível em outras línguas, inclusive em línguas geneticamente próximas do PB. Em francês, por exemplo, a tradução literal de 26) é totalmente inaceitável. Mas uma possível tradução seria 30), onde a alteração da expressão sintática dos argumentos do verbo é assinalada por uma construção causativa marcada pelo verbo *faire*, acompanhada do clítico *se*, e conseqüentemente do auxiliar *être*:

- 30) Jean s'est fait opérer.
 Jean SE Aux Caus operar

O PE, por sua vez, é mais exigente do que o PB, mas menos do que o francês. Nessa língua, o pronome *se* é suficiente para apontar para o fato de que *João* não é o agente da operação, mas seu paciente. Mas crucialmente, uma frase como 26) não pode ter esse sentido. Voltamos então ao que já observamos acima: 26, para um locutor de PB, é ambígua, podendo ser associadas às duas estruturas subjacentes seguintes:

- 26') O João_i operou cv_i

26'') O João_i operou cv_j

Em 26', o sujeito e o objeto do verbo são interpretados como a mesma pessoa, enquanto em 26''), como duas pessoas distintas. Já para um locutor de PE, só a segunda interpretação é possível.

Em resumo, tudo que foi dito acima nos leva a caracterizar o PB como uma língua tendo um funcionamento sintático bastante distinto do PE, e, aliás, das línguas românicas em geral. Esse funcionamento foi inicialmente caracterizado como sendo de língua orientada para o tópico. Ou seja, o tópico no PB é uma noção integrada à gramática de frase. Mais importante ainda é que isso não significa simplesmente que achamos em PB enunciados impossíveis em PE. Isso é verdadeiro, mas tem uma implicação mais profunda se trabalharmos com a noção de Língua-I. Significa que as estruturas associadas aos enunciados, *inclusive a enunciados idênticos*, pelos falantes de PB, são diferentes das estruturas associadas pelos falantes de PE. Em outros termos, a diferença não se limita aos enunciados diferentes, mas perpassa toda a língua.

Na próxima seção, aprofundaremos a questão da diferença gramatical entre as línguas, introduzindo a noção de parâmetro.

III. O igual e o diferente: gramática universal e gramáticas particulares

Enfatizamos nas duas seções anteriores a relevância de se distinguirem os enunciados, ou Língua-Externa, e a gramática, ou Língua-Interna. A gramática, desse ponto de vista, é entendida como a competência que permite que os falantes associem estruturas aos enunciados. A comparação do PE e do PB nos mostrou que enunciados superficialmente iguais podem corresponder a estruturas diferentes associadas a eles por gramáticas diferentes.

Toda a argumentação acima foi no sentido de mostrar que duas línguas aparentemente muito parecidas podem corresponder a gramáticas distintas. Mas é possível inverter o paradoxo e mostrar que duas línguas aparentemente muito diferentes não o são tanto assim e compartilham propriedades fundamentais. A Teoria de Princípios e Parâmetros (Chomsky 1981, 1985, 1995) procura formular, por um lado, os princípios gerais da Gramática Universal, supostamente compartilhada por todas as línguas naturais, e, por outro lado, os parâmetros de variação que explicam as diferenças entre elas. Segundo Baker (2001):

“Um parâmetro é simplesmente um ponto de escolha na receita geral das línguas humanas. Um parâmetro é um ingrediente que pode ser acrescentado para fazer um tipo de língua, ou deixado de lado para fazer um outro tipo. Um parâmetro pode também ser um procedimento de combinação que pode ser feito de duas ou três maneiras para dar dois ou três tipos de línguas diferentes. Se você toma os ingredientes genéricos da linguagem, acrescenta o tempero B e chacoalha, você obtém o inglês. Se você toma os mesmos ingredientes básicos, mas em lugar do tempero B, você acrescenta o aditivo D e E e agita, você obtém o navajo. As Línguas-I são receitas e os parâmetros, os poucos passos básicos nessas receitas onde as diferenças entre as línguas podem ser criadas.” (op. cit. p. 57)

Os princípios constituem assim a receita geral das línguas humanas, com que cada criança vem equipada ao nascer. Algumas escolhas, porém, ficam em aberto, que no processo de aquisição, expostas a uma realização particular da linguagem humana, as crianças terão que fazer.

Quais são os ingredientes básicos da linguagem? Mencionamos, de maneira genérica, os princípios da gramática universal. Mas precisamos especificar primeiro os objetos sobre os quais esses princípios operam. As categorias como Nome, Verbo, Adjetivo, etc.²² podem ser consideradas como primitivos da linguagem. Desde os primórdios da tradição gramatical, tentou-se propôr uma lista exaustiva daquilo que se

²² Para Diomedo, as partes do discurso são oito: nome, pronome, verbo, particípio, advérbio, conjunção, preposição, interjeição. Os leitores curiosos poderão facilmente perceber que essa lista ainda está muito atual nas gramáticas escolares.

costumava chamar as “partes do discurso”. A lingüística moderna retomou muito da tradição, inovando, porém, num aspecto. Na tradição gramatical que nos vem dos antigos, partes do discurso correspondem a palavras, uma vez que, como vimos na introdução, “o discurso é uma composição de palavras”. Diomedo também define a palavra (*dictio*) como um som articulado provido de sentido. Ora, o estruturalismo apontou para o fato que a menor unidade correspondendo a essa definição não era a palavra, mas o morfema. Este é o assunto da seção V, mas é importante, desde já, apontar para o fato de que os primitivos da gramática não correspondem forçosamente a palavras. O fato de uma determinada categoria poder ser realizada ou não como uma palavra independente - ou seja, como morfema “livre” – é, aliás, uma das importantes diferenças entre as línguas. O artigo, por exemplo, que é expresso por um morfema livre em português, é incorporado ao nome, como uma marca flexional, em outras línguas. Em línguas de morfologia muito rica, como as línguas chamadas polissintéticas, uma frase inteira pode ser expressa por uma única palavra, composta de uma grande quantidade de morfemas expressando todas as relações envolvidas na frase²³. Essas categorias, freqüentemente realizadas como morfemas presos, passaram a ser alvo de atenção especial na Teoria da Gramática Gerativa nos anos 80, ao serem consideradas como essenciais na geração das orações e o *locus* essencial das diferenças entre línguas. Passou-se assim a distinguir as categorias “lexicais”, como Nome e Verbo, das categorias “funcionais”, como Tempo, Concordância, Determinante, Complementador e outras que chegaram a ser definidas na literatura. As primeiras correspondem a listas grandes e abertas, sujeitas a permanentes inovações, de itens lexicais portadores de traços semânticos complexos. As segundas, ao contrário, são expressas por conjuntos limitados de itens lexicais, resistentes a inovações, e, geralmente, caracterizáveis por traços binários (+/-). O sufixo ‘-ar’ do português, por exemplo, pode ser caracterizado correspondendo à categoria Tempo, com o traço [-finito].

. Novos verbos do português podem ser criados acrescentando à raiz verbal esse sufixo.

²³ Cf. Sândalo (2001) e Seção V.

Vejam-se, por exemplo, os neologismos “bitnetar” ou “lular”, nascidos nos anos 90 no Brasil, e já em desuso. Nenhuma inovação lexical, contudo, consiste em substituir o sufixo ‘-ar’ por outra realização morfológica. Esses elementos são extremamente resistentes a mudanças, e quando essas acontecem, elas envolvem profundas alterações na gramática como um todo.²⁴

Apesar de terem propriedades claramente distintas, as categorias funcionais e lexicais obedecem aos mesmos princípios de construção, ou projeção, da frase. Da mesma maneira que um nome projeta obrigatoriamente um sintagma nominal (ou NP, do inglês *Noun Phrase*) e um verbo um sintagma verbal (ou VP, do inglês *Verb Phrase*) quando entram numa derivação sintática, a categoria Tempo projeta um TP (do inglês *Tense Phrase*), a categoria Concordância, um AgrP (do inglês *Agreement Phrase*), a categoria Determinante, um DP (do inglês *Determiner Phrase*), a categoria Complementador, um CP (do inglês *Complementizer Phrase*), etc.... Além disso, a estrutura interna dos NPs, VPs, TPs, DPs, e CPs, é regida pelos mesmos princípios. Em particular, pelo *Princípio de Endocentricidade* que diz que todo sintagma deve ter um núcleo de mesma categoria. Assim como todo NP tem um núcleo N, todo TP tem núcleo T. É visto no sentido oposto, o correlato da projeção referida acima: todo núcleo X^0 projeta XP, e todo XP tem X^0 como núcleo. Assim se pode dizer que XP é a projeção máxima de X^0 ²⁵. Adicionalmente, as relações definidas estruturalmente no interior dos sintagmas são as mesmas para todas as categorias e podem ser representadas da seguinte maneira:

$$31) \quad [{}_{XP} \textit{Especificador} [{}_X \textit{X Complemento}]]$$

²⁴ Temos um exemplo de mudança na morfologia de tempo na história do inglês, onde o sufixo de infinitivo do inglês antigo foi substituído pela preposição *to*. Sabemos que essa mudança se acompanhou numa grande reorganização da sintaxe do verbo na língua.

²⁵ O conjunto de princípios que rege o formato dos sintagmas é conhecido como Teoria X' (leia-se Teoria X-barra), para uma apresentação mais detalhada e formal, ver Raposo (1992) e Miotto et al. (2004), Cap. II.

31) representa dois níveis de projeção para toda categoria. O primeiro (em negrito) contem o núcleo e seu complemento. O segundo contem essa primeira projeção e o especificador do núcleo. O especificador pode ser definido como a categoria que entra numa relação de concordância com traços do núcleo²⁶.

Os princípios gerais que regem a construção de sintagmas a partir de categorias são parte da receita geral para as línguas naturais, ou seja, a gramática universal. O que varia de língua para língua são “detalhes” como a ordem na qual núcleo e complementos ou núcleo e especificadores se encontram. Uma língua como o lakhota, por exemplo, citada por Baker (2002) instancia uma imagem simetricamente oposta à do inglês, ou do português, na ordem linear dos complementos em relação aos núcleos:

32) John [found [that letter] [under [the bed]]]

João achou essa carta debaixo da cama.

33) [John [[[wowapi k'uhe] [oyuke ki] ohlate] iyeye]

John letter that bed the under found

John carta essa cama a debaixo achou

Comparando 32) e 33) vemos que em todos os sintagmas a ordem entre o núcleo e o complemento é inversa: o verbo precede em inglês e segue em lakhota todos os seus complementos, a preposição do inglês e do português é uma posposição em lakhota, os determinantes precedem os nomes em inglês e português e os seguem em lakhota.

²⁶ Note-se que, para constituir um sintagma bem formado, só a presença do núcleo é requerida (pelo Princípio de Endocentricidade). Assim, a presença de complemento e de especificador não é obrigatória, mas depende das propriedades do núcleo. Decorre disso que as noções de projeção máxima e intermediária são relativas. Assim, se um núcleo não tem especificador numa determinada estrutura, a categoria que contem esse núcleo e seu complemento se torna a sua projeção máxima. O caso extremo é aquele em que um sintagma só contem o núcleo. Isso é freqüente acontecer com sintagmas nominais só compostos por um nome, como *maçã* em *Gosto de maçã*.

A aparente grande diferença entre essas línguas pode se reduzir então a uma pequena diferença na receita: os complementos seguem os núcleos em português e inglês, mas precedem-nos em lakota.

Combinando as diversas possibilidades, podemos derivar as diversas ordens de palavras nas línguas do mundo, cuja primeira descrição sistemática se deve ao lingüista tipologista Joseph Greenberg.

Os outros grandes princípios que constituem a receita básica das línguas humanas dizem respeito às restrições sobre o estabelecimento de relações entre elementos da estrutura. Apresentamos acima a Teoria de ligação que rege as relações referenciais entre expressões nominais. Destacamos a atuação, nessa teoria, da *Condição de c-comando* que restringe a configuração estrutural na qual se define a relação de ligação (cf. nota 10). Como já mencionamos, a *Condição de c-comando* atua também na definição de uma outra relação sintática fundamental, a regência, que desempenha um papel essencial tanto no deslocamento dos elementos nas estruturas, quanto na atribuição do caso abstrato. Independentemente da realização morfológica observável em certas línguas - mas não em todas - o caso pode ser considerado como um elemento essencial na visibilidade da funções semânticas instanciadas na frase. Em frases como:

- 34) Leo venatorem occidit.
- 35) O leão matou o caçador.

sabemos que, em latim, foi *o leão* que matou *o caçador*, e não o contrário, porque *leo* está no caso nominativo, e *venatorem* no caso acusativo e, em português, porque *o leão* precede o sujeito e *o caçador* o segue. Se invertermos a ordem linear no português e a marcação morfológica dos casos em latim, obteremos a interpretação oposta. Dentro do raciocínio desenvolvido aqui, o latim e o português compartilham uma propriedade essencial: o caso abstrato obrigatoriamente atribuído aos sintagmas nominais. Essa

propriedade faz parte da receita geral. A maneira como esse caso é expresso é o ingrediente que produz a diferença entre as duas línguas. Em latim, a expressão é morfológica, em português, é configuracional. Línguas como o português requerem configurações estruturais nas quais os atribuidores de caso regem o sintagma nominal que recebe o caso. Como já repetimos várias vezes, regência envolve c-comando. Ou seja, uma categoria A rege uma categoria B se, e somente se, A c-comanda B. Mas não é suficiente. Uma outra propriedade das línguas naturais, presente na teoria da ligação, também emerge na noção de regência: a localidade. Como também já vimos anteriormente, as relações sintáticas são locais, ou seja, elas se dão num determinado domínio. A localidade também afeta o movimento. Por exemplo, não se pode mover uma palavra interrogativa de dentro de orações também interrogativas. É a razão pela qual, na frase 36), *quando* não pode ser interpretado como *quando Maria foi embora*, mas somente como *quando João disse*.

36) Quando que o João disse por que a Maria foi embora?

Admitindo que o deslocamento de *quando* deixa uma categoria vazia, a representação da interpretação impossível da frase seria:

36') Quando_i que o João disse porque_j a Maria foi embora cv_i?

36) contrasta com a frase 37) na qual *quando* pode ser interpretado tanto como modificando *pensou*, quanto modificando *foi embora*.

37) Quando_i que o João pensou que a Maria iria embora cv_i?

A diferença entre 36) e 37) é que, na primeira, existe um elemento interrogativo na oração subordinada que cria um domínio para a interpretação da categoria vazia. Já em 37), na ausência de tal elemento, a frase fica ambígua.²⁷ Vale notar que esse fenômeno se encontra inclusive nas línguas que não tem movimento aparente dos sintagmas interrogativos, como o chinês. Em chinês, a ordem linear do equivalente de 36) é 38):

38) O João disse por que a Maria foi embora quando?

O fato de essa frase ser mal formada em chinês, apesar de *quando* estar na posição em que seria interpretado, pode ser explicado se postularmos que a diferença entre o chinês e o português está na visibilidade do movimento do elemento interrogativo. Desse ponto de vista, os interrogativos se movem em todas as línguas para a posição inicial da frase (possivelmente porque o traço interrogativo que domina a frase como um todo provoca esse movimento) e em todas as línguas a interpretação envolve uma categoria vazia. Mas a diferença entre as línguas é que, em algumas, esse movimento é visível e em outras não é²⁸.

Uma última propriedade universal das línguas que vale mencionar é que quando um item lexical é projetado numa derivação sintática, as categorias que ele seleciona semanticamente ou sintaticamente devem ser projetadas também. No caso da seleção semântica, e em particular em relação aos verbos, falaremos de estrutura argumental. Esse princípio geral explica porque frases como 39) e 40) em inglês são mal formadas:

39) * John put on the table yesterday.

²⁷ O fato de uma frase ser ambígua não impede que uma das interpretações seja preferida pelos falantes, por razões que têm a ver com o processamento. Assim, fora de contexto, 37) será mais facilmente interpretada com “quando” modificando “pensou”.

²⁸ Uma maneira de dar conta dessa diferença é postular dois lugares diferentes para o movimento na gramática, tal como é representada pelo modelo em T. Se ele se dá antes da Estrutura-S, ele é acessível à interpretação fonológica e, portanto, visível. Se ele se dá depois da Estrutura-S, no componente que associa a Estrutura-S à Forma Lógica, ele não é acessível à interpretação fonológica e, portanto, invisível.

John pôs na mesa ontem.

40) * Put the book on the table yesterday.

Pôs o livro na mesa ontem.

O verbo *to put* é um verbo transitivo, que seleciona dois argumentos, um argumento externo, seu sujeito, e um argumento interno, seu objeto. As frases do inglês são mal formadas porque, em 39), o objeto está ausente e, em 40), o sujeito está ausente. O leitor atento terá notado que as traduções para o português correspondem a frases bem formadas. O português é uma língua na qual tanto o sujeito quanto o objeto podem ser omitidos. A propriedade do sujeito nulo se encontra em muitas línguas do mundo. Um exemplo bastante estudado são as línguas românicas, menos o francês. A propriedade do objeto nulo parece menos frequente. Nas línguas românicas, o português - tanto PE quanto PB - é a única onde 39) é aceitável. Mas o PE e o PB diferem na extensão do fenômeno (cf. Galves 2001, cap. 8). Em relação a esse fenômeno, o PB se parece de novo muito com o chinês, que permite tanto sujeito, quanto objeto nulo. Assumindo o raciocínio desenvolvido acima a respeito dos lugares de diferença das línguas, podemos pensar que o que difere entre o PB e o inglês não é a aplicação do princípio de projeção dos argumentos na sintaxe, mas antes, a gramaticalidade de uma estrutura em que a posição sujeito e objeto são ocupadas por uma categoria vazia como representado em 39' e 40':

39' John pôs *cvi* na mesa ontem

40' *cvi* pôs o livro na mesa ontem

Essas duas estruturas respeitam o princípio universal da projeção, uma vez que as posições sujeito e objeto existem - apesar de não serem visíveis. O que diferencia a gramática do PB da gramática do inglês, é que ela permite a identificação referencial das

categorias vazias nessas posições, ou seja, permite que elas sejam interpretadas como remetendo a entidades definidas no discurso. Note-se novamente que o recurso ao discurso para a interpretação das categorias vazias é mediatizado pela sintaxe, uma vez que, em inglês, esse recurso não é disponível. Podemos pensar, por tudo o que foi dito acima, que, em PB, a identificação referencial das categorias vazias sujeito e objeto faz parte do conjunto de propriedades associadas à caracterização da língua como “língua orientada para o tópico”. No PE, que mostramos não ter o mesmo funcionamento, o sujeito nulo é legitimado pela flexão rica no verbo, como nas outras línguas românicas de sujeito nulo²⁹.

IV. Variação e mudança

Afirmamos acima que as diferentes línguas são como uma variação sobre um mesmo tema. Alguns ingredientes particulares, os parâmetros, são capazes de dar uma aparência muito diferente à faculdade de linguagem comum a todos os humanos. A variação porém, não é só encontrada quando se comparam línguas diferentes, ela aparece também no interior das línguas. Nesta seção, veremos como uma teoria de princípios e parâmetros, como a que foi apresentada acima, trata a variação intralingüística.

Outras correntes teóricas argumentam que a variação é inerente à língua. É o caso da chamada Teoria da Variação, cujo principal expoente é o lingüista americano William Labov. Para ele, a gramática não é de natureza determinística, como no modelo gerativo, mas probabilístico. Várias formas competem para a mesma função e fatores de natureza lingüística ou social favorecem a realização de uma ou outra forma. Com isso, a língua

²⁹ A identificação referencial do objeto nulo no PE é um assunto mais complexo que não desenvolveremos aqui. Para leituras complementares ver Raposo (1986), Duarte (1987), Cyrino (1993, 1996) e Galves (1987, 1988, 1998, e 2001, cap. 8).

também está em permanente mudança e a distinção entre diacronia e sincronia, enfatizada por Saussure no seu curso de *Linguística Geral*, se desfaz.

A Teoria da Variação foi inicialmente desenvolvida no âmbito de estudos fonológicos, mas teve um impacto importante também na sintaxe³⁰. Coerentemente com a concepção de linguagem referida acima, os estudos variacionistas em sintaxe se caracterizam por dois aspectos: na descrição, a metodologia quantitativa, e na interpretação, a concepção funcionalista da linguagem. Porém, a associação sistemática da primeira com a segunda foi colocada em questão por Tarallo e Kato (1989), que propuseram articular à metodologia quantitativa o conceito chomskiano de parâmetro, criando uma corrente nova intitulada “Socio-linguística paramétrica”. Esta associação, bastante frutífera, renovou o olhar sobre os fenômenos de variação sintática no português brasileiro ao trazer uma abordagem mais dedutiva na descrição dos mesmos. Partindo de uma teoria dotada de uma capacidade de predição forte, em que um só parâmetro de variação abrange uma gama grande de fenômenos, alguns dos quais aparentemente distintos entre si e onde os estudos sobre uma determinada língua podem ser projetados sobre outras línguas aparentemente muito diferentes, a “socio-linguística paramétrica” permitiu uma compreensão mais abrangente da relação entre variações correlatas. Porém, além dos avanços descritivos, uma questão epistemológica permanece: não será contraditório aliar uma apreensão quantitativa dos fatos sintáticos associada a uma concepção da linguagem em que tudo é variação a uma teoria em que a gramática é um objeto determinístico? Como resolver esse paradoxo sem perder o avanço que a consideração da variação pode trazer aos estudos gramaticais e o aumento de poder explicativo que a teoria da gramática pode trazer ao estudo da variação?

Do ponto de vista da Teoria da Gramática Gerativa, a distinção Língua-Interna/Língua-Externa nos permite, de fato, articular de maneira não contraditória gramática e variação. A variação é um fenômeno de Língua-E. Se circunscrevemos o

³⁰ Cf. Paiva e Scherre (1999).

nosso objeto de estudo à Língua-I, ela fica excluída dos estudos de cunho gerativista. É efetivamente o que ocorreu na área durante muito tempo. Porém, nos anos 90, com o desenvolvimento dos estudos diacrônicos, a questão da variação e da mudança voltou a integrar a pauta dos estudos gerativistas e, com ela, a relação na mudança da Língua-I e da Língua-E. Essa relação é de mão dupla. Por um lado, para retomar as palavras de Pintzuck, Soulas e Warner (2000), o desafio é “interpretar a Língua-E dos textos históricos em relação à Língua-I dos seus falantes (op. cit. p.). Por outro lado, temos o desafio colocado por Givón:

“Todas as pressões funcional-adaptativas que dão forma à estrutura sincrônica – idealizada – da linguagem são exercidas durante o efetivo desempenho. É onde a linguagem é adquirida e onde a gramática emerge e muda. É onde a forma se ajusta – criativamente e sob a impulsão da construção oportunística momentânea do contexto – a novas funções e sentidos estendidos. É também onde variação e indeterminação são os ingredientes necessários do mecanismo efetivo que modela e remodela a competência.”(op. cit. p. 6)

É importante notar que, neste trecho, o funcionalista Givón põe em relação aquisição e mudança, que, para os gerativistas, são termos intimamente associados, já que do seu ponto de vista, a gramática muda na aquisição quando crianças de uma determinada geração escolhem uma gramática diferente da geração anterior. O modelo é claro. A questão de difícil elucidação, porém, é como se dá o ajuste entre os dados, variáveis e indeterminados, e a modelagem da competência. Essa questão tem sido recentemente posta na agenda dos gerativistas (cf. Kroch 2001 e Lightfoot 1991, 1999, entre outros). Remetemos o leitor interessado ao capítulo sobre mudança lingüística para mais detalhes.

Uma última questão, essencial, resta ser elucidada. Se a gramática é determinística, qual é a origem da variação? A resposta é dupla. Por um lado, as gramáticas produzem variação de natureza sintática. Uma língua de sujeito nulo, ou de objeto nulo, por exemplo, também pode ter sujeito ou objeto pronominal exposto. Uma

língua de ordem Sujeito-Verbo pode produzir, de maneira mais ou menos abrangente, a ordem Verbo-Sujeito dependendo da construção ou do próprio verbo. Uma língua de ordem Objeto-Verbo permite a ordem Verbo-Objeto em construções de extraposição do objeto. Uma língua V2, em que o verbo aparece normalmente na segunda posição, pode produzir, ainda que de maneira restritiva, ordens V3. A lista é longa e dificilmente exaustiva.

Por outro lado, as comunidades lingüísticas não são forçosamente homogêneas do ponto de vista gramatical. Isso é particularmente verdadeiro em períodos de contato entre populações, por razões externas à própria língua: contatos, imigrações, invasões, etc... Esses períodos são geralmente geradores de mudança lingüística e podem ser traduzidos lingüisticamente em termos de competição de gramáticas (cf. Kroch 2001). Deve se notar também que um tipo particular de competição de gramáticas é gerador nos dados de uma quantidade importante de variação: a competição entre a gramática do vernacular e a gramática, ou gramáticas, ou fragmentos de gramáticas veiculados pela norma. Esta tem dois importantes veículos nas sociedades modernas: a língua escrita e a escola. Essa competição é claramente visível numa língua como o português brasileiro, onde a norma veiculada pela escola ainda contém muitos resquícios dos estágios passados da língua. Um caso que chama particularmente a atenção é o da utilização e colocação dos pronomes clíticos. O clítico de terceira pessoa e a colocação enclítica são claramente adquiridos na escola (cf. Correa 1992), eles são também objeto de sistemática revisão nos livros e jornais. A sua utilização na língua falada tende a desaparecer hoje, mas esse processo não se completou e ele é afetado por fatores de natureza sócio-cultural.

Do ponto de vista da gramática gerativa, essa variação deve ser distinguida da outra. Cabe ao lingüista fazer essa distinção, tarefa nem sempre fácil. O ponto que é essencial enfatizar é que ambas são compatíveis com uma concepção de linguagem em que a gramática é um objeto determinístico.

V. De volta à palavra

Como já foi mencionado no início deste capítulo, a ‘palavra’ há muito vem sendo abordada como tema central por estudiosos da linguagem, seja através da tentativa de exprimir uma definição para ela, seja através da tentativa de uma descrição de sua composição estrutural. Mas quando se toma como objeto de estudo lingüístico ‘a palavra’ em uma dada língua, defini-la parece não ser uma tarefa tão simples.

Retomando os conceitos de palavra advindos da gramática tradicional e apresentados no início deste capítulo, temos as seguintes definições: *‘uma PALAVRA é constituída de elementos materiais (vogais, consoantes, semivogais, sílabas, acento tônico) a que se dá um sentido e que se presta a uma classificação’*; *‘palavra é uma unidade lingüística de som e significado que entra na composição dos enunciados da língua’*.

De posse destas definições de palavra, em português, é possível classificarmos tanto ‘amaremos’ como ‘-re’ ou ‘-mos’ como palavras. Isto porque tanto ‘amaremos’ quanto ‘-re’ ou ‘-mos’ são sons constituídos de elementos materiais como consoantes e vogais, carregam um sentido ou significado (‘amaremos’ = primeira pessoa do plural do futuro do presente do verbo ‘amar’ que, por sua vez, carrega o sentido de ‘ter o sentimento que predispõe alguém a desejar o bem de outrem’; ‘-re’ = carrega o significado de tempo futuro do presente; ‘-mos’ = carrega o significado de concordância de primeira pessoa do plural) e entram na composição dos enunciados da língua. Todavia, pelo nosso conhecimento de falantes de português, sabemos que ‘amaremos’ é uma palavra e ‘-re’ e ‘-mos’, assim como ‘am-’ (raiz verbal) e ‘-a-’ (vogal temática), são

elementos componentes da palavra ‘amaremos’. Estes elementos componentes são denominados *morfemas*, unidades mínimas de significação da Morfologia.

Em línguas isolantes como o chinês, cada palavra carrega apenas um significado, mas em línguas polissintéticas como o kadiwéu, falada no Mato Grosso do Sul, certas seqüências de sons, assumidas por seus falantes como palavras, carregam significados traduzidos por frases em línguas como o português (cf. Sândalo, 2001). Por exemplo, segundo Sândalo, a seqüência de sons ‘*jotaγanγetaγadomitiwaji*’ do kadiwéu significa, em português, ‘eu falo com eles por vocês’. Como, então, saber se ‘*jotaγanγetaγadomitiwaji*’ é uma palavra ou uma frase em kadiwéu?

Observe que as definições de palavra – ‘som articulado com uma significação’; ‘elemento formado por unidades menores, os morfemas’ – e frase – ‘enunciação de sentido completo’ – advindas das gramáticas tradicionais não nos ajudam muito a saber se ‘*jotaγanγetaγadomitiwaji*’ é uma palavra ou uma frase em kadiwéu. Por estas definições, ‘*jotaγanγetaγadomitiwaji*’ pode ser tanto um som articulado com uma significação e um elemento formado pela composição de unidades menores (*j* = sujeito de primeira pessoa + *otaγan-γen*: = falar – transitivizador + *-t-γa* = - ? –objeto indireto de segunda pessoa + *-dom* = benefactivo + *i* = plural do objeto indireto + *-t-waji* = -?- plural do objeto direto), quanto uma enunciação de sentido completo, conforme a tradução para o português: ‘eu falo com eles por vocês’.

A literatura fonológica lingüística nos traz a noção de ‘palavra prosódica’, que pode ser definida como um elemento contendo um único acento principal³¹. Por esta definição, ‘amaremos’ é uma palavra prosódica porque contém um único acento principal na sílaba ‘re’ e ‘am-’, ‘-a-’, ‘-re’ e ‘-mos’, isoladamente, não são palavras prosódicas

³¹ Para definição de ‘palavra prosódica’, conferir, entre outros, Nespor e Vogel (1986) e Vigário (2003).

porque não contêm um acento principal. Por esta mesma definição, em português brasileiro, em uma sentença como *Lhe deram o livro*, nas expressões *o livro*, ou *lhe deram*, o artigo *o* e o pronome clítico *lhe*, isoladamente, não podem ser considerados palavras prosódicas porque são elementos átonos. Assim, *o livro* é uma única palavra prosódica, bem como *lhe deram*, uma vez que *o livro* porta um único acento principal (sílabas ‘li’) e *lhe deram* porta um único acento principal (sílabas ‘de’). A definição de palavra prosódica ainda não é suficiente para definirmos uma palavra em português, posto que *o*, assim como *lhe*, são considerados palavras diferentes de *livro* e *deram* nesta mesma língua.

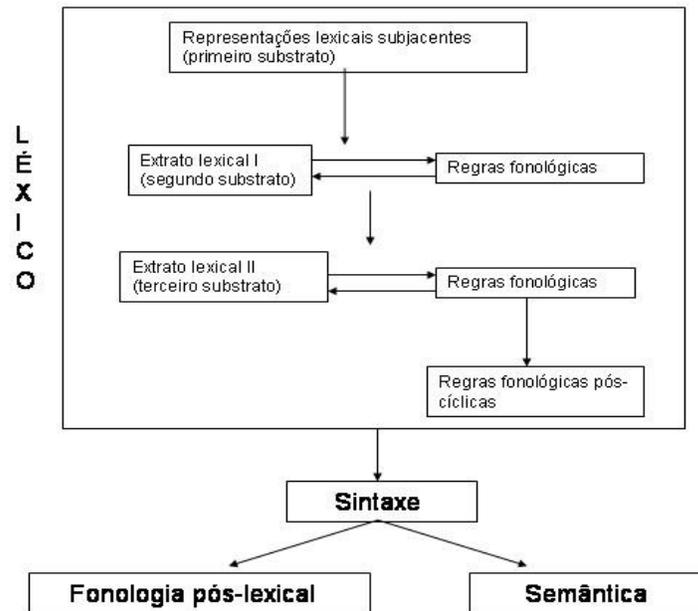
Assim, de uma maneira simplificada, mas suficiente para nossos propósitos, consideraremos aqui que a ‘palavra’ pode ser definida como um elemento formado por morfemas, que possuem significado na estrutura interna das palavras, mas que não podem aparecer isoladamente carregando ‘significação’ em um enunciado. Já as palavras carregam significação própria dentro do enunciado. Por esta definição, isoladamente no interior de um enunciado como *Lhe deram o livro* apresentado acima, *livro* é uma palavra porque carrega o significado ‘obra literária científica ou artística que compõe em regra, um volume’, *o* é uma palavra porque carrega o sentido de ‘definitude’ e gênero masculino que será atribuído a algum nome, *lhe* é uma palavra porque carrega o sentido de ‘a ele’, *deram* é uma palavra porque carrega o sentido de ‘eles exerceram o ato dar (doar, conceder)’. Da mesma forma, em um enunciado em português como *Amaremos nossos filhos*, *amaremos* é uma palavra porque carrega o significado ‘nós exerceremos o ato de amar (desejar o bem)’. Porém ‘-re’ e ‘-mos’ só podem expressar, respectivamente,

o significado de ‘tempo futuro’ e de concordância de ‘primeira pessoa do plural’ no interior de ‘amaremos’ e não isoladamente no enunciado mencionado.

Dado que a palavra é formada por morfemas, resta-nos ainda saber onde e como se dá a formação das palavras, ou seja, onde e como se dá a união dos morfemas na formação das palavras na arquitetura da gramática das línguas.

Nas seções anteriores deste capítulo, foi discutida a gramática da frase (domínio do componente sintático) na abordagem gerativista. Assumindo que a morfologia é o componente da gramática que trata da palavra e de seus componentes (os morfemas), podemos nos perguntar onde está a morfologia na arquitetura da gramática. Se adotarmos o modelo em T de arquitetura de gramática (cf. seção I.4 deste capítulo), podemos pensar que a morfologia está no léxico, onde haveria operações morfológicas sendo realizadas independentemente dos diversos componentes da gramática, conforme a representação a seguir.

Este tipo de arquitetura de gramática, que é assumido pelas teorias gerativas das décadas de 70 e 80, pressupõe que cada componente da gramática deve corresponder a um módulo independente, governado por seus princípios particulares. Portanto, cada componente da gramática (fonologia, sintaxe, semântica) é tratado de modo independente dos outros componentes. O componente fonológico é dividido em dois níveis: a Fonologia Lexical (processada no léxico) e a Fonologia Pós-Lexical (processada após a derivação sintática). Uma vez que a Morfologia apresenta uma forte relação com a Fonologia, a Morfologia é parte da Fonologia Lexical (localizada no léxico). Eis a representação da arquitetura da gramática que descrevemos:



Neste tipo de arquitetura, o léxico é constituído de substratos onde os morfemas são adicionados uns aos outros na formação das palavras e onde as regras fonológicas são aplicadas depois da adição de cada morfema.

No primeiro substrato estão as representações lexicais subjacentes, como lista de morfemas, condições de estrutura morfológica e formas irregulares. O segundo substrato lexical (extrato lexical I) é responsável pela aplicação de regras de formação de palavras por derivação, pela adição dos morfemas derivacionais. Os morfemas derivacionais possuem a característica de alterar a categoria gramatical de uma palavra, não são produtivos, ou seja, não é qualquer morfema derivacional que pode ser adicionado a qualquer raiz e há muitas restrições de co-ocorrência operando sobre eles. Um exemplo de morfema derivacional é o morfema ‘*izar*’ que transforma substantivos em verbos em português:

41) hospital → hospitalizar;

Nota-se que enquanto o morfema ‘izar’ é adicionado após o substantivo *hospital*, o mesmo morfema não pode ser adicionado após o substantivo *hóspede* (*hospedizar). Deve-se dizer *hospedar* e é preciso memorizar que é possível a produção de *hospitalizar*, mas não de **hospedizar*. Este tipo de argumento é utilizado para reforçar a idéia de que o léxico é um local de armazenamento de irregularidades.

Já o terceiro substrato (extrato lexical II) também é responsável pela aplicação de regras de formação de palavras, porém, não por derivação como no extrato lexical I, mas por flexão, através do acréscimo dos morfemas flexionais, como por exemplo, acréscimo dos morfemas de plural aos nomes e dos morfemas de flexão de tempo e pessoa aos verbos. Os morfemas flexionais não alteram categorias de palavras como os morfemas derivacionais, mas estabelecem relações entre as palavras, como relações de concordância. Eles são produtivos: por exemplo, qualquer verbo pode ser marcado por um morfema indicando primeira pessoa do plural (em português: cantaremos, leremos, iremos) e qualquer artigo pode ser pluralizado (em português: as, os, uns, umas). Exceções são mais raras no paradigma flexional e mais freqüentes no paradigma derivacional. Como ilustração de formação de palavras por flexão, trazemos a formação da primeira pessoa do plural no futuro do presente do verbo ‘cantar’ em português:

42) cant (raiz verbal) + a (vogal temática) + re (sufixo de tempo futuro do presente) + mos (sufixo de concordância de pessoa: primeira pessoa do plural) = cantaremos (primeira pessoa do plural do verbo *cantar* no futuro do presente)

Após a aplicação das regras no extrato lexical I e no extrato lexical II, podem ocorrer as regras fonológicas pós-cíclicas ainda no componente lexical, como as regras de harmonia vocálica. Por exemplo, em português, a vogal ‘e’ da palavra ‘meninos’ pode sofrer alçamento neste contexto em que é seguida pela vogal alta ‘i’ da sílaba subsequente da mesma palavra:

43) meninos → [mininus]

Após o componente lexical, há o componente sintático que é responsável pela derivação sintática, ou seja, é o componente responsável, entre outras coisas, pela formação de sintagmas e pelos movimentos de núcleos e de projeções máximas. E, por sua vez, após o componente sintático, há o componente fonológico, no qual regras pós-lexicais podem ocorrer, como as regras de reajuste na cadeia segmental.³²

Conforme o que foi exposto, para os modelos morfológicos gerativos desenvolvidos nas décadas de 70 e 80, a derivação das palavras ocorre de um modo independente dos outros componentes da gramática. O léxico é um local de armazenamento de irregularidades memorizadas e como cada língua tem seu léxico específico, fica explicada a diversidade encontrada nos domínios da morfologia. Conforme Chomsky (1970), toda e qualquer operação lexical precede a sintaxe. Esta

³² Como exemplo de regra de reajuste na cadeia segmental citamos o exemplo de sândi entre palavras. Abaixo, temos, como exemplo, o sândi vocálico representado pelo apagamento da vogal ‘a’ no final da palavra ‘casa’ quando seguida pela mesma vogal no início da palavra ‘azul’ que a segue:

(i) casa azul → [cazazu]

manipula palavras inteiras e, conseqüentemente, é cega à estrutura interna das palavras, ou seja, às operações lexicais. Esta visão ficou sendo conhecida como hipótese lexicalista.

Porém, certas ocorrências morfológicas, como os morfemas flexionais já explicitados anteriormente, trazem evidência de que a morfologia não parece atuar unicamente em um léxico independente dos diversos componentes da gramática e de que a Sintaxe não pode ser cega à Morfologia (cf. Anderson, 1982).

Como vimos na Seção III, em línguas como o latim, a ocorrência dos morfemas flexionais de caso em núcleos de sintagmas nominais indica o papel sintático destes sintagmas. A presença do caso nominativo indica que dado sintagma nominal é sujeito, assim como a presença do caso acusativo indica que dado sintagma nominal é objeto direto e a presença do caso dativo indica que dado sintagma nominal é objeto indireto:

- 44) Puer magistram basiat.
menino professora beija
'O menino beija a professora'.

No exemplo acima, 'puer' está no caso nominativo e é sujeito na sentença latina, assim como a palavra 'magistram' que está no caso acusativo, o que é percebido pela partícula '-m', é objeto direto na mesma sentença. Isto parece indicar que a Morfologia não pode ser toda processada em um léxico independente e anterior à Sintaxe, posto que morfemas flexionais, como os morfemas de caso, remetem a estruturas sintáticas.

Os morfemas de flexão verbal, por sua vez, trazem evidência de que a Sintaxe não pode ser cega às operações morfológicas. No exemplo (42), os morfemas ‘-re’ e ‘-mos’ do verbo *cantaremos*, são, respectivamente, morfemas flexionais de tempo (futuro do presente) e pessoa (primeira pessoa do plural) em português. Em línguas como o português, que apresentam flexão verbal, os morfemas flexionais presentes nos verbos são alojados em projeções sintáticas específicas na derivação do componente sintático como projeções de tempo e concordância (respectivamente, T e Agr), conforme o que foi apresentado a respeito da projeção das categorias funcionais na seção III deste mesmo capítulo. Uma consequência imediata para estas línguas é a operação sintática de movimento do verbo para estas projeções, o que ilustra o fato de que a Sintaxe não pode ser cega à Morfologia.³³

Outro exemplo da forte interação entre a Morfologia e os outros componentes da gramática é encontrado em Baker (1988). Baker mostra que na língua chichewa o morfema aplicativo ‘-ir’ que se afixa ao verbo no léxico, uma vez que não pode aparecer como constituinte independente em qualquer nível sintático, pode aparecer na realização fonética como ‘-ir’ ou ‘-er’, dependendo do contexto fonológico. O que determina a distribuição destas duas formas é uma regra geral de harmonia vocálica. Segundo esta regra, a vogal tensa /i/ aparece depois de radicais verbais cuja última vogal é tensa (/i/, /u/ ou /a/); por sua vez, a vogal não-tensa /e/ aparece depois de radicais verbais cuja última vogal é não tensa (/e/ ou /o/):

³³ No entanto, cabe notar que, assumindo uma perspectiva gerativa minimalista, ainda é possível considerar um sistema gramatical no qual as palavras são formadas no léxico e as operações sintáticas se dão por motivos de checagem dos traços morfológicos (cf. Chomsky, 1995). Assumindo esta perspectiva, no exemplo em português, o verbo *cantaremos* é alçado para T e Agr para checar traços de tempo e concordância. Neste tipo de abordagem a Sintaxe também não seria cega à Morfologia.

- 45) Msangalatsi a- ku- yend- **er-** a ndodo.
artista sp-pres.-andar-aplic.-asp bastão.
'O artista anda com um bastão.'
- 46) Mbalame zi- ma- uluk- **ir-** a mapiko.
pássaros sp- hab- voar- aplic.- asp asas.
'Os pássaros voam com (usando) asas'.

O morfema aplicativo 'ir/er' expressa a função 'instrumental', função esta que, em muitas línguas, é expressa por um advérbio ou um adjunto adverbial realizado por um sintagma preposicional. Nas respectivas traduções dos exemplos em chichewa para o português, a função de instrumental expressada pelo morfema 'ir/er' naquela língua, é expressa, em português, através dos sintagmas preposicionais 'com um bastão' e 'com asas'.

Baker afirma que o mesmo princípio morfofonológico que determina os modelos de combinações formados no léxico, determina também os modelos de combinações formadas na Sintaxe .

Assim como os exemplos do morfema de caso no latim e dos morfemas de flexão verbal de tempo e pessoa em português, o exemplo do morfema aplicativo '-ir' em chichewa parece nos revelar a dificuldade em conceber um sistema no qual a Morfologia aparece exclusivamente no léxico e atuando de maneira completamente independente dos outros componentes da gramática.

Portanto, uma vez que a Morfologia não pode estar única e independentemente associada ao léxico, ainda nos resta saber a resposta da questão colocada no início desta seção: Onde estaria a Morfologia na arquitetura da gramática?

Levando em conta todas as considerações já feitas, as possíveis alternativas para a abordagem do componente morfológico no sistema gramatical gerativo podem ser, entre outras: (i) considerar um sistema gramatical no qual a morfologia é uma subteoria que perpassa todo o sistema e que atua segundo um conjunto de princípios de boa-formação de palavras e; (ii) considerar que um único sistema gramatical gerativo é capaz de gerar tanto frases quanto palavras e no qual a morfologia se encontra ‘distribuída’ nos diversos componentes da gramática. Estas duas abordagens são encontradas respectivamente em Baker (1988) e na Morfologia Distribuída (doravante MD).

Baker constrói um modelo no qual a morfologia não é um componente da gramática, mas uma subteoria, ou seja, um sistema de princípios, como a Teoria da Ligação, ou a Teoria X’. Segundo ele, a ‘teoria morfológica’ pode ser caracterizada como uma teoria que se aplica quando uma estrutura complexa de forma [${}_Z$ X + Y] é criada, ou seja, uma teoria que se aplica quando estruturas são formadas no nível nuclear. Para Baker, o mais importante efeito da teoria morfológica na Sintaxe é filtrar certas formações impossíveis nas línguas. Por exemplo, restrições morfológicas bloqueiam sintagmas no interior de uma palavra. Desta maneira, em português, se um dado sintagma complemento, como *a bola*, for inserido no interior de um verbo, como *pegou*, o resultado é uma formação agramatical como: **pegabolaou*. Esta condição morfológica de boa-formação, que bloqueia a criação de componentes impossíveis no léxico, também impedirá algumas estruturas de serem formadas na Sintaxe. Isto explica por que, na Sintaxe, adjunções a um elemento nuclear são possíveis para outros elementos nucleares, mas não para projeções máximas.

Assim, diferentemente dos modelos que consideram as operações morfológicas como parte do próprio léxico, o ponto de vista que considera a morfologia como um sistema de princípios, além de capturar restrições gramaticais, captura também generalizações morfofonológicas de uma forma integrada no sistema gramatical.

Mais recentemente, uma das teorias morfológicas, desenvolvida no âmbito do quadro gerativo, e que vai ao encontro das noções presentes em Baker (1988) sobre a relevância da Morfologia para a derivação ocorrida nos outros componentes gramaticais e sobre a semelhança da derivação na Sintaxe com a derivação na Morfologia é a Morfologia Distribuída (doravante MD, cf. Halle & Marantz (1993 e 1994)).

Para a MD, a estrutura morfológica é sintática. Na arquitetura de gramática proposta por esta teoria, palavras e sintagmas são formados por um único sistema gerativo. Este sistema possui um conjunto de regras que geram estruturas sintáticas sujeitas a operações morfológicas na derivação da forma fonológica (PF).

Diferentemente dos modelos gerativos das décadas de 70 e 80, na MD não há a idéia de um léxico independente dos componentes sintático e fonológico e governado por seus princípios particulares onde estão contidas todas as informações lexicais. Nesta teoria, as informações lexicais estão agrupadas em diferentes listas conforme o tipo e se encontram distribuídas através dos componentes da gramática. Estas listas são de três tipos: lista de terminais sintáticos, vocabulário e enciclopédia.

A lista de terminais sintáticos é a primeira lista acessada na derivação da gramática. Ela se encontra no léxico e é acessada antes da derivação sintática. Os terminais sintáticos são os *nós terminais* das estruturas sintáticas (os núcleos) e correspondem a morfemas. Cada morfema é um complexo de traços que podem ser de

dois tipos: fonológicos ou sintático-semânticos. Dependendo dos traços que constituem os morfemas eles serão definidos como *raízes* ou *morfemas abstratos*. A distinção entre raízes e morfemas abstratos corresponde, de um modo geral, à distinção estabelecida acima (cf. seção III deste mesmo capítulo) entre categorias lexicais e funcionais (cf. Harley e Noyer, 1998).

Os *morfemas abstratos* correspondem aos itens pertencentes a categorias funcionais e são compostos exclusivamente de traços não- fonéticos, tais como [passado], [plural], ou traços que compõem o determinante D.

As *raízes* correspondem a itens pertencentes à categoria lexical e são constituídas por traços fonéticos. As raízes não contêm traços sintáticos, mas apenas traços fonéticos e conteúdo semântico. Como exemplos de raízes podemos citar, em inglês: \sqrt{dog} (cachorro), \sqrt{ox} (boi), \sqrt{mouse} (rato), etc.

A segunda lista acessada na derivação gramatical é o vocabulário. Este é uma lista de itens vocabulares que é acessada na forma fonológica (PF), onde há a aplicação da operação de *inserção vocabular* que provê material fonológico aos morfemas abstratos.

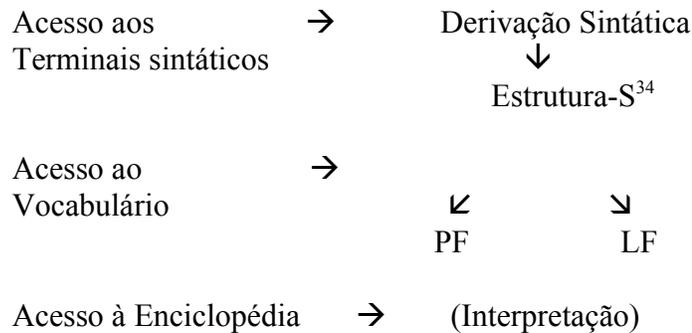
A inserção vocabular é responsável, por exemplo, pela atribuição de material fonológico ao morfema abstrato de plural. Em inglês, os diferentes materiais fonológicos que expressam o plural nos nomes – ‘z’, ‘en’, ‘Ø’ – são atribuídos ao morfema abstrato de plural conforme princípios que regulam a inserção vocabular. Assim, temos a atribuição do material fonológico ‘z’ ao morfema abstrato de plural no contexto de presença da raiz \sqrt{dog} (cachorro), a atribuição do material fonológico ‘en’ ao morfema abstrato de plural no contexto de presença da raiz \sqrt{ox} (boi) e a atribuição do material

fonológico ‘ \emptyset ’ ao morfema abstrato de plural no contexto de presença da raiz \sqrt{mouse} (rato):

- 47) plural de *dog* = *dog*[z];
 plural de *ox* = *oxen*
 plural de *mouse* = *mice*.

A terceira e última lista acessada na derivação é a enciclopédia. Ela é uma lista que contém informações semânticas que devem ser listadas como uma propriedade de uma raiz ou de um objeto construído sintaticamente, como as expressões idiomáticas. Por exemplo, na enciclopédia, há a informação semântica de que, em português, a expressão idiomática *quebrar o galho* significa *ajudar*. A enciclopédia é acessada posteriormente à saída da forma fonológica (PF) e da forma lógica (LF), ou seja, na interpretação.

Eis a representação do acesso às listas na derivação da gramática:



³⁴ Na morfologia distribuída, que se baseia na arquitetura da gramática proposta no âmbito do Programa Minimalista (cf. Chomsky, 1995), não há propriamente Estrutura-S, mas uma operação chamada *Spell-out* (literalmente “pronuncia”). Contrariamente à Estrutura-S, *Spell-Out* não é um nível de representação mas uma operação que efetua a ligação da estrutura derivada até esse ponto com o componente PF.

Nesse modelo, portanto, além da morfologia interagir fortemente com os outros componentes da gramática, uma vez que as operações morfológicas se aplicam de maneira ‘distribuída’ no sistema gramatical, palavras e sintagmas são gerados por um único sistema gerativo. Ou seja, o mesmo sistema gerativo que produz frases é responsável pela realização sintática (formação dos sintagmas e movimentos destes) e morfológica (formação de palavras) destas frases.

Por exemplo, em uma frase como *Pedro chutou a bola*, o mesmo sistema gramatical gerativo é responsável: (i) pela formação do sintagma DP *Pedro*, bem como pelo seu movimento sintático para o especificador de TP e pelo acesso à raiz ; (ii) pela formação do sintagma VP *chutou*, pelo acesso à raiz e ao morfema abstrato que contém o traço de tempo passado, pela inserção de material fonológico ‘ou’ a este morfema e pela operação de movimento que agrega o morfema ‘-ou’ à raiz e ainda; (iii) pela formação do sintagma DP ‘a bola’, pelo acesso à raiz e ao morfema abstrato que contém os traços [+definido], [+feminino] e [+singular] no nó D, assim como pela inserção do material fonológico ‘a’ a este morfema.

Isso aproxima fortemente as noções de palavra e frase, dado que tanto uma palavra quanto uma frase são geradas de maneira semelhante por um único sistema gerativo.

VI. Considerações finais

O nosso ponto de partida neste capítulo foram as definições tradicionais associadas a palavra e frase, em termos da relação entre forma e sentido. Procuramos mostrar como esta relação é problematizada na lingüística moderna. Argumentamos que para dar conta da interpretação semântica da frase, convem definir dois níveis de regras, as que pertencem à gramática de frase e as que pertencem ao discurso. Os princípios que regem as primeiras são os mesmos que regem processos sintáticos. Crucialmente, esses princípios são sensíveis à estrutura. Além disso, eles podem fechar certas opções, apontando para a precedência lógica da gramática sobre o discurso.

Mostramos também a relevância de se distinguir Língua-Interna e Língua-Externa como dois objetos distintos. A primeira é um objeto abstrato, uma representação da competência dos falantes nativos. Argumentamos que só nesse nível é que se pode discutir a semelhança ou diferença do PB e do PE de um ponto de vista genuinamente lingüístico. Enfatizamos a necessidade de postular elementos abstratos nas representações das sentenças, que têm uma interpretação semântica, mas carecem de realização morfológica: as chamadas “categorias vazias”. Discutimos o lugar da diferença entre as línguas numa teoria que postula que os princípios gerais da gramática são universais.

Grande parte da discussão do capítulo foi dedicada à questão da relação entre os diversos componentes e níveis de representação na gramática. Apontamos para a relação privilegiada entre o léxico e a sintaxe e tematizamos o lugar da morfologia na arquitetura geral da gramática. Voltando ao fim à questão inicial da palavra e da frase, fechamos o capítulo apresentando teorias recentes que aproximam ao máximo essas duas noções pré-teóricas, ao afirmar a identidade dos processos de construção de uma e de outra.

Referências bibliográficas

Abaurre, M. L.; Pontara, M. N.; Fadel, T. (2003) *Português: língua e literatura: volume único (2a. Edição)*. São Paulo: Moderna.

Anderson, S. (1982) “Where is Morphology?” *Linguistic Inquiry*, 13, pp. 571-612.

Baker, M. (1988) *Incorporation: a theory of grammatical function changing*, Chicago: Chicago University Press.

_____ (2001) *The atoms of language, the mind’s hidden rules of grammar*, Basic Books.

Uma belíssima e original introdução à teoria de Parâmetros.

Cegalla, D. P. (1998) *Novíssima gramática da língua portuguesa*. São Paulo: Editora Nacional.

Charpin, F. (1980) “L’héritage de l’antiquité dans la terminologie grammaticale moderne”, *Langue Française*, 47, pp. 25-32.

Chomsky, N. (1964) *Aspects of the Theory of Syntax*, trad. Portuguesa Eduardo Raposo, *Aspectos da teoria sintática*, Lisboa, 1975.

Livro fundador da Teoria Padrão da Gramática Gerativa

_____ (1970) “Remarks on Nominalization”, in R. Jacobs & P. Rosenbaum (orgs.) *Readings in English transformational grammar*, Waltham, MA: Ginn.

_____ (1972) *Studies on Semantics in Generative Grammar*, The Hague: Mouton.

_____ (1975) *Reflections on Language*, New York, Pantheon.

Traz o modelo da Teoria Padrão estendida (que inclui a teoria dos vestígios, o embrião da teoria de princípios, e a teoria lexicalista), e uma discussão com a Semântica Gerativa e a Teoria dos Atos de fala

_____ (1981) *Lectures on Government and Binding*, Dordrecht, Foris.

Livro fundador da “Teoria de Regência e Ligação” rebatizada depois “Teoria de Princípios e Parâmetros”

_____ (1985) *Knowledge of Language, its nature, origin and use*, Praeger, trad. port. Inês Duarte (org.) *Conhecimento da Língua, sua natureza, origem e uso*, Lisboa, Caminho, 1994.

_____ (1995) *The Minimalist program*, trad. port. Eduardo Raposo, *O programa minimalista*, Lisboa, Caminho, 1999.

Cinque, G (1983) “Topic-constructions in some European languages and connectedness”, em K. Erlich e H. Van Riemsdijk (orgs.) *Connectedness in sentence, discourse and text. Tilburg Studies in Language and Litterature*, vol.4.

Corrêa, V. R. (1992). O objeto direto nulo no português do Brasil, .Dissertação de mestrado, Universidade Estadual de Campinas.

Cunha, C. (1972) *Gramática do Português Contemporâneo*. Belo Horizonte: Editora Bernardo Álvares S. A.

Duarte, I. (1987) A construção de topicalização na gramática do português: regência, ligação e condições sobre o movimento, Tese de doutorado inédita, Universidade de Lisboa.

Duarte, M.E. (1995) A perda do princípio Evite Pronome no português brasileiro, Tese de doutorado inédita, Universidade de Campinas.

Galves, C. (1987) “A sintaxe do português brasileiro”, *Ensaio de Linguística*, 13, Belo Horizonte, pp. 31-50.

_____ (1988) "Objeto nulo e predicação : hipóteses para uma caracterização do português brasileiro", *D.E.L.T.A.* , 4.2, São Paulo, 1988, pp.273-290

_____ (1998a) “A gramática do português brasileiro”, *Línguas e instrumentos lingüísticos*, 1, pp. 79-96.

_____ (1998b) “Tópicos, sujeitos, pronomes e concordância no português brasileiro”, *Cadernos de Estudos Lingüísticos*, 34, pp. 7-21.

_____ (2001) *Ensaio sobre as gramáticas do português*, Editora da Unicamp.

Givón, T. (2001) *Syntax*, John Benjamins.

Greenberg, J. (1963) *Universals of Language*, Cambridge: MIT Press.

Halle, M. e Marantz, A. (1993) “Distributed Morphology and the Pieces of Inflection.”, Em K. Hale e S. Keyser (eds.) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 111-176.

_____ (1994). “Some key features of Distributed Morphology”. Em, A. Carnie e H. Harley. (orgs) *MITWPL 21: Papers on phonology and morphology* Cambridge, pp. 275-288.

Harley, H. e Noyer, R. (1998) “Licensing in the non-lexicalist lexicon: nominalizations, vocabulary items and the Encyclopaedia”. H. Harley. *MITWPL 32: Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, Cambridge, pp. 119-137.

Kato, M. (1993) “Recontando a história das relatives em uma perspectiva paramétrica”, in I. Roberts e M. Kato (orgs.) *Português brasileiro, uma viagem diacrônica*, Campinas: Editora da Unicamp, pp. 223-261.

Kato, M. e Ramos, J. (1999) “Trinta anos de sintaxe gerativa no Brasil, *D.E.L.T.A.*, vol. 15, no especial: *Trinta anos de Abralín*, pp.105-146.

Retrospectiva dos estudos gerativistas sobre a sintaxe do português brasileiro

Kidway, A. (1999) “Word Order and Focus Positions in Universal Grammar”. In: G. Rebuschi and L. Tuller (eds). *The Grammar of Focus*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kroch, A. (1994) “Morphosyntactic Variation” *Proceedings of the 30th annual meeting of the Chicago Linguistics Society*, vol 2, pp. 180-201.

_____ (2001) “Syntactic Change”, in M. Baltin e C. Collins (orgs.), *Handbook of Syntax*, Blackwell.

Lightfoot, D. (1991) *How to set parameters: arguments from Language Change*, MIT Press.

_____ (1999) *The development of language, acquisition, change and evolution*, Blackwell.

Labov, W. (1975) *Sociolinguistic patterns*, Philadelphia: University of Pennsylvania Press.

Li, C. e S. Thompson (1976) “Subject and topic, a new typology for language”, in C. Li (org.) *Subject and Topic*. Academic Press.

Lobato, L. (1986) *Sintaxe gerativa do português: da teoria padrão à teoria de regência e ligação*, Belo Horizonte: Vigília.

Marantz, A. (1984) *On the nature of grammatical relations*. Cambridge: MIT Press.

Mioto, C. Figueiredo Silva, M. C. e Lopes, R. (2004) *Novo manual de sintaxe*, Florianópolis, Editora Insular.

Negrão, Esmeralda (1999) O Português brasileiro: Uma língua voltada para o discurso, Tese inédita de Livre-Docência, Universidade de São Paulo.

Nespor, M. e Vogel, I.(1986) *Prosodic phonology*. Dordrecht: Foris Publications.

Paiva, M.C. e Scherre, M. (1999) “Retrospectiva socio-linguística: contribuições do PEUL”, , *D.E.L.T.A*, vol. 15, no especial: *Trinta anos de Abralín*, pp. 201-232.

Pintzuck, S., Soulas, G. e Warner, A. (orgs.) (2000) *Diachronic Syntax, Models and Mechanisms*, Oxford University Press.

Pontes, E. (1981) Da importância do tópico em português, *Anais do 5º Encontro de Linguística*, PUC-RJ; republicado em *O tópico no português do Brasil*, Pontes 1987.

Trabalho pioneiro sobre a caracterização do português como língua de tópico.

Raposo, E. (1992) *Teoria da Gramática. A Faculdade de Linguagem*, Caminhos, Lisboa.

Livro clássico de introdução, em português, ao modelo de princípios e parâmetros da Teoria da Gramática Gerativa.

Sândalo, M.F. (2001) “Morfologia”, in F. Mussalim e A.C. Bentes (orgs.) *Introdução à Linguística: domínios e fronteiras*, vol.1, São Paulo: Cortez.

Tarallo F. e M. Kato (1989) Harmonia trans-sistêmica: variação intra e intra-linguística, Preedicação 5, UNICAMP.

O texto fundador da “Socio-linguística paramétrica”.

Vigário, M. (2003) *The prosodic word in European Portuguese*, Mouton de Gruyter.