

A conceituação clássica do morfema

3.1. Introdução

Vimos no capítulo anterior que o início do século xx assistiu ao questionamento e subsequente declínio de importância da noção de *palavra*. Neste capítulo vamos apresentar a noção de morfema, na sua versão estruturalista norte-americana, bem como algumas das críticas feitas a essa proposta. Tais críticas prepararam o retorno da noção de palavra à morfologia.

A *palavra* havia sido o fundamento da gramática tradicional. Mas como definir essa unidade? Despojada da representação escrita – vista como “meramente um dispositivo externo” (Bloomfield, 1933: 294) que reproduzia imperfeitamente a fala de uma comunidade (*id.*: 293) –, a delimitação da palavra tornava-se difícil. Não coincidia, na maioria das vezes, com um elemento mínimo de som e significado, e sua característica distintiva passava a ser a possibilidade de ser enunciada em isolado. Nada de muito interessante.

Os problemas com a noção de palavra apontados pelos estruturalistas decorriam, em grande parte, de a definirem como uma **forma**, *i.e.*, como “um traço vocal recorrente que tem significado” (Bloomfield, 1926: 27). Isto implicava haver a necessidade da utilização de critérios fonológicos indissociados de critérios gramaticais para a sua apreensão. Fonologicamente uma sequência como *deixe-me*, por exemplo, é uma palavra, uma vez que *me* equivale a uma sílaba átona em relação ao verbo e não pode, sozinho, funcionar como enunciado. Gramaticalmente, porém, *deixe-me* equivale a duas palavras: *me* é um pronome em função de objeto e pode ser mudado de posição para antes do verbo, o que não acontece com simples sílabas.

Para evitar que enunciados diferentes pudessem ser segmentados de maneiras diversas e que noções oriundas dos estudos tradicionais fossem associadas à análise gramatical, a lingüística do século xx retirou da noção de *palavra*, em favor da noção de *morfema*, a ênfase que tinha nos séculos anteriores. O morfema tornou-se a unidade básica da gramática e, por conseguinte, da morfologia – agora transformada em **morfologia baseada em morfemas**. Desse modo, a morfologia da maior parte do século xx passou a ser a análise sintagmática dos vocábulos¹.

¹ Estamos utilizando indistintamente os termos *palavra* e *vocábulo*.

3.2. Três modelos de análise lingüística

Tal mudança correspondeu à adoção de um modelo de análise gramatical diferente daquele herdado da tradição greco-latina. O estruturalismo norte-americano estabeleceu um método para identificar que partes específicas do material fonológico de uma forma complexa expressavam as diferentes partes de um significado também complexo. As unidades som e significado assim depreendidas eram os **elementos mínimos** ou **itens** da análise.

3.2.1. A palavra no centro da análise: o modelo *Palavra e Paradigma*

Para ilustrar a diferença entre o modelo de análise que prevalecera por tantos séculos no Ocidente – denominado por Hockett (1954) **Palavra e Paradigma** (ou *PP*) – e o modelo de análise estruturalista norte-americano dominante entre meados da década de 1940 e de 1950² – que ficou conhecido como **Item e Arranjo** (ou *IA*) –, tome-se como exemplo o estudo tradicional do verbo em português.

Um verbo como AMAR, por exemplo, é regular e pertence à primeira conjugação. Qualquer das formas desse verbo é descrita pelas suas características de *palavra morfossintática* (v. cap. 5). A referência a *amemos*, por exemplo, é feita como em (3.1):

(3.1) primeira pessoa do plural do presente do subjuntivo de AMAR

Qual a diferença entre essa e uma análise estruturalista? A diferença está em que a caracterização em (3.1) não é linear (Matthews, 1972: 106): *amemos* é um todo, não a relação sintagmática entre signos mínimos que se devem suceder necessariamente numa dada ordem, como expresso a seguir:

(3.2) AM- + VT/ SUB PRES + 1PL

Exemplificado do modo clássico em (3.3) adiante, o paradigma apresenta as diversas realizações da palavra, ou *lexema*, que no nosso exemplo é AMAR. Tais formas resultam das variações ao longo de cada *categoria gramatical* admitida pelo verbo em português e das combinações entre as diferentes categorias. Cada categoria – Número, por exemplo – concretiza-se como *propriedades morfossintáticas* (como Singular, Plural) que se opõem e excluem.

No modelo tradicional, a relação entre a cadeia sonora e as propriedades expressas não é direta. Não se trata de uma sucessão de formas mínimas, como se os dados pertencessem sempre a línguas aglutinantes ideais (até porque esse modelo foi desenvolvido no estudo de línguas que exibem morfologia flexional razoavelmente complexa). Não se procurava, no modelo gramatical que nos foi legado pela tradição greco-

² Nem todos os estruturalistas norte-americanos seguiram tão de perto IA. O próprio Bloomfield estava bem mais orientado para IP.

latina, decompor palavras em unidades mínimas de som e significado, em busca da seqüência sonora específica que corresponde a tal ou qual significado e vice-versa. Aliás, a função do hífen nas formas em (3.3) adiante não é de assinalar a segmentação em unidades mínimas de som e significado, mas a de indicar que qualquer verbo regular de primeira conjugação pode seguir esse **paradigma** (a palavra grega para 'modelo'), que funciona, descritivamente, como uma espécie de molde, ou, se preferirmos, como uma fronteira para regras que derivam formas do paradigma. Para isso substitui-se: (a) ou o elemento na posição inicial, (*am-*) – o **radical**, –, por outro de mesma classe, como *louv-*, por exemplo, e mantêm-se as **terminações**, que expressam as categorias gramaticais envolvidas; ou (b) mantêm-se o radical e mudam-se as terminações. As relações assim expressas são *verticais*, ou melhor, paradigmáticas.

Para ficar mais evidente o quanto a análise tradicional estava afastada de qualquer preocupação com formas mínimas, relembremos, por um momento, o modo como aprendemos a conjugar o verbo em latim no colégio, seguindo um modelo tão antigo quanto o gramático Prisciano (ca. 500 d.C.) – daí a denominação **formação prisciânica** para o estudo da formação do paradigma segundo esse modelo.

O verbo latino que em português traduzimos por 'optar, desejar' deveria ser lembrado por quatro formas: *optō*, *-āvi*, *-āre*, *-ātum*. Tais formas correspondiam, respectivamente:

- a) à **primeira pessoa do presente do indicativo ativo** (*optō*);
- b) à **primeira pessoa do singular do perfeito do indicativo ativo** (*optāvī*);
- c) ao **infinitivo presente ativo** (*optāre*);
- d) ao **supino**³ (*optatum*).

A partir dessas formas – que representam os chamados *tempos primitivos* – podia-se conjugar todo o verbo. Vamos usar aqui para ilustração apenas as formas ativas. Do **primeiro radical** (ou **raiz de formação**) formavam-se o *imperfeito* e o *futuro do indicativo*, o *presente do subjuntivo*, o *particípio presente* e o *gerúndio*. Bastava para isso trocar a terminação: tirando-se o *-ō* final de *optō* e colocando-se *-abam*⁴ tinha-se o *imperfeito* (*optabam*); se em vez de *-abam* a terminação escolhida fosse *-abo*, tinha-se o *futuro* (*optabo*); para o *gerúndio*, tirava-se o *-ō* final e colocava-se *-andi*, *-ando* ou *-andum* (*optandi*, *-o*, *-um*).

Do **segundo radical** formavam-se o *mais-que-perfeito* e o *futuro perfeito do indicativo*, o *perfeito* e o *mais-que-perfeito do subjuntivo* e o *infinitivo perfeito*, trocando-se o *-ī* final, respectivamente, por *-eram* (*optaveram*), *-ero* (*optavero*), *-erim* (*optaverim*), *-issem* (*optavissent*), *-isse* (*optavisse*). Do **terceiro radical** formavam-se

³ Forma nominal do verbo latino em *-um*, empregada com verbos de movimento, ou em *-u*, com adjetivos. O supino pode ser traduzido como indicando a finalidade (algo como *lusum*, 'para jogar', *auditum* 'para ouvir').

⁴ Estamos exemplificando sempre com a primeira pessoa do singular. O mecanismo era o mesmo para as demais pessoas.

dois tempos: pela supressão da terminação *-re*, o *imperativo (opta)*; do acréscimo de *-m*, o *imperfecto do subjuntivo (optarem)*. Por fim, do *supino* em *-um* formava-se o *participio futuro*, pela troca de *-um* por *-urus (optaturus, -a, -um)*.

Em resumo: derivamos *palavras* do paradigma verbal a partir de outras *palavras* do mesmo paradigma. Não partimos de seqüências de formas mínimas mais abstratas para explicar quer a formação dos chamados tempos primitivos quer a dos demais tempos. Como não estamos trabalhando com signos mínimos, também não vem ao caso termos formado, por exemplo, o *imperfecto* e o *futuro* a partir do *presente*. Para um estruturalista, tudo isso é muito estranho, para dizer-se o mínimo.

(3.3) AMAR

Indicativo

PRESENTE	PRETÉRITO PERFEITO	PRETÉRITO IMPERFEITO
1SG am-o	am-ei	am-ava
2SG am-as	am-aste	am-avas
3SG am-a	am-ou	am-ava
1PL am-amos	am-amos	am-ávamos
2PL am-ais	am-astes	am-áveis
3PL am-am	am-aram	am-avam

PRETÉRITO MAIS-QUE-PERFEITO	FUTURO DO PRESENTE	FUTURO DO PRETÉRITO
1SG am-ara	am-arei	am-aria
2SG am-aras	am-arás	am-arias
3SG am-ara	am-ará	am-aria
1PL am-áramos	am-aremos	am-aríamos
2PL am-áreis	am-areis	am-aríeis
3PL am-aram	am-arão	am-ariam

Subjuntivo

PRESENTE	IMPERFEITO	FUTURO
1SG am-e	am-asse	am-ar
2SG am-es	am-asses	am-ares
3SG am-e	am-asse	am-ar
1PL am-emos	am-ássemos	am-amos
2PL am-eis	am-ásseis	am-ardes
3PL am-em	am-assem	am-arem

Imperativo

AFIRMATIVO	NEGATIVO
am-a	Não am-es
am-e	Não am-e
am-emos	Não am-emos
am-ai	Não am-eis
am-em	Não am-em

Formas nominais

INFINITIVO IMPESSOAL	INFINITIVO PESSOAL	GERÚNDIO	PARTICÍPIO
am-ar	am-ar	am-ando	am-ado
	am-ares		
	am-ar		
	am-amos		
	am-ardes		
	am-arem		

3.2.2. O morfema no centro da análise:
o modelo *Item e Arranjo*

Para um estruturalista, quaisquer das formas latinas que acabamos de ver são seqüências de unidades mínimas de som e significado que se ajustam a um **padrão geral** – o equivalente descritivo dos *paradigmas dos verbos regulares* dos estudos tradicionais – que descreve a flexão verbal na língua em estudo. Não há a necessidade de um paradigma para que de uma palavra se derive outra. Uma forma como *optaturus*, numa análise estrutural, é resultante da seqüência em (3.4):

(3.4) opt-	ã -	t -	ur	-us
desejar	VT	marca do	PART. FUT	NOM/SG
		4° radical		

É uma análise sintagmática. O papel da morfologia passa a ser o de juntar esses pedaços de material, ou *morfemas*, que são signos mínimos. Para identificá-los no enunciado, o estruturalismo lançou mão da **substituição** ou **teste da comutação**⁵, que se apoiava no “princípio primordial da gramática” (Camara Jr., 1973: 43): a *oposição lingüística*. Retomando-se o verbo AMAR como exemplo, uma forma como *amas* contém uma parte que não se altera, que é *am-*, que se relaciona ao significado mais geral

⁵ *Substituição* é termo proposto por Harris e utilizado na lingüística norte-americana; *teste da comutação* é a denominação dada por Hjelmslev e seguida, no Brasil, por Camara Jr. (v. Haugen, 1951). Para críticas, ver o quarto capítulo de Chomsky (1964).

atribuído a AMAR. Substituído *am-* por *louv-*, ou por *ador-*, por exemplo, tem-se uma forma que apresenta um outro *significado lexical*, o que nos faz concluir que *am-* é uma unidade de som e significado; um morfema, portanto.

Como saber, porém, se o restante, *-as*, é um elemento ou mais de um? Também pelo confronto com outras formas. *Amas* difere, por exemplo: (a) da forma de *terceira pessoa do singular (ama)*, o que nos mostra que *-s* pode ser isolado; (b) da forma do *subjuntivo (ames)*, o que demonstra que *-a-* também pode ser isolado; por fim, (c) não há qualquer marca específica para Tempo-Modo-Aspecto (doravante TMA), como no *futuro (amarás)*, por exemplo, ou no *pretérito imperfeito (amavas)*, formas, por essa razão, mais “canônicas” do que *amas*. *Amas* poderia ser compreendida, portanto, numa análise em constituintes imediatos (ou *CIs*) como ou contendo um zero (3.5a), ou ainda, seguindo Pontes (1965)⁶, como em (3.5b), contendo um *morfe cumulativo*⁷:

(3.5)

a.			
amas			
am-	-a-	- ø -	-s
raiz	VT	IND. PRES.	2SG

b.		
amas		
am-	-a-	-s
raiz	VT + IND. PRES	2SG

Uma análise do verbo em português deveria, pois, arrolar quais os *elementos* ou *itens* constitutivos de sua estrutura, como também os padrões em que se apresentam. Daí denominar-se esse tipo de análise *Item e Arranjo* ou *Elemento e Arranjo*. Cada um dos elementos mínimos pertence a uma *ordem* ou *classe* (raiz, afixo). Elementos da mesma ordem são mutuamente exclusivos para uma única posição (Nida, 1949: 84): isto significa que uma palavra não pode ao mesmo tempo estar no Futuro e no Passado, por exemplo, numa língua que expresse essas propriedades.

3.2.3. Uma tradição do Oriente: *Item e Processo*

Um terceiro modelo de análise, mais antigo que IA, porque suas origens remontam aos trabalhos de gramáticos hindus sobre o sânscrito, como Pāṇini (ca. 500 a.C.),

⁶ Pontes (1965: 60): “Consideramos que a VT, no Presente do Indicativo, indica também MTA, porque existe a oposição entre este paradigma e todos os outros. Não se pode analisar como um zero morfológico, neste caso, porque o sufixo MTA também não pode ser segmentado no Pretérito Perfeito, que, no entanto, se opõe (com exceção de P3) ao Presente do Indicativo. A interpretação como morfema cumulativo se impõe”.

⁷ Ver nota 22 adiante.

foi denominado **Item e Processo** (ou **IP**). Tal como IA, IP é um modelo aglutinativo (v. Spencer, 1991: 50), embora a concatenação de elementos se dê num nível mais abstrato que aquele do enunciado, o das formas **subjacentes** ou **teóricas**. A essas formas são aplicados *processos*, ou *regras*, ou *operações*, que as transformam nas formas de superfície⁸. Esse artifício procura demonstrar a regularidade que existe por trás das irregularidades, que seriam, então, irregularidades aparentes.

Tomemos para exemplo a descrição proposta por Camara Jr. (1970: 104) para as formas verbais *hás, há, hão*. Segundo o A., elas derivam das estruturas subjacentes em (3.6). O asterisco nas formas da primeira coluna indica que se trata de formas teóricas, que não se realizam como tais fonologicamente:

(3.6)	*havs	⇒	haø	⇒	hás
	*hav	⇒	haø	⇒	há
	*hav/N/	⇒	haø/N/	⇒	hão

Por não ser seguido da vogal temática da segunda conjugação, o radical **hav-* perde a consoante final /v/ (perda que assinalamos com o símbolo ø) e forma sílaba com a consoante seguinte, no caso de ela existir.

Os mesmos dados levam a análises diferentes, se nosso modelo é IA: como uma sucessão de zeros morfológicos, ou, ainda, com a postulação de um *morfe cumulativo*. A forma *há*, por exemplo, seria a cumulação de “tema, MTA e PN” (Pontes, 1965: 73-74), isto é, “[’a | P2, Presente do Indicativo, verbo haver” (*id.*: 74)⁹.

A distinção entre os três modelos é fundamental para entendermos a definição clássica de morfema e alguns dos problemas que esse novo elemento apresentou aos pesquisadores, questões de que trataremos a seguir.

3.3. A definição clássica de morfema¹⁰

O **morfema** é “uma forma recorrente (com significado) que não pode ser analisada em formas recorrentes (significativas) menores” (Bloomfield, 1926: 27); “uma forma lingüística que não mantém semelhança fonético-semântica com qualquer outra forma” (Bloomfield, 1933: 161). Estas definições de Bloomfield:

... dão conta do morfema em duas dimensões, e correspondem a dois passos fundamentais no

⁸ O formato que a análise toma em IP é o que se chama *derivação*: camadas estruturais que se formam sucessivamente, pela aplicação de operações a uma determinada cadeia de elementos. *Derivação*, nesse sentido, não tem o significado mais comum nos estudos lingüísticos, quando se opõe a *flexão*.

⁹ Ver, porém, a seção 3.5.4. adiante.

¹⁰ Lembramos que o quadro a que nos reportamos é o do estruturalismo norte-americano, especialmente àquele formado pelos distribucionalistas das décadas de 1940 e 1950 que ficaram conhecidos como bloomfieldianos.

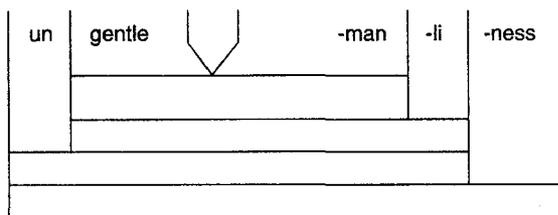
... dão conta do morfema em duas dimensões, e correspondem a dois passos fundamentais no processo de identificação de morfemas. No primeiro, o da *segmentação*, serão isoladas na cadeia da fala seqüências fônicas recorrentes mínimas que apresentam significado; no segundo, o da *classificação*, serão considerados membros do mesmo morfema os morfemas que apresentem disintitvidade fonético-semântica comum... (Basílio, 1974a: 80 – ênfase adicionada).

Cada morfema é um átomo de som e significado – isto é, um signo mínimo. Segundo tal perspectiva, a morfologia é o estudo desses átomos (a *alomorfia*) e das combinações em que podem ocorrer (a *morfotática*) – i.e., a **morfologia** é o estudo dos *morfemas e de seus arranjos*¹¹.

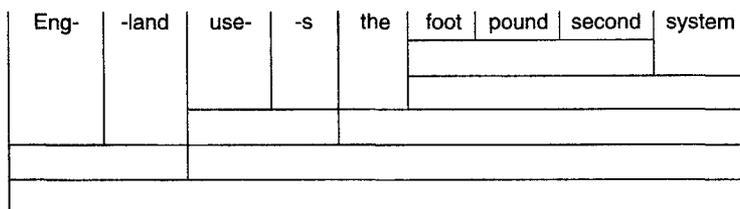
Com Bloomfield o *signo* deixava de ser a palavra (como fora para Saussure) e passava a ser, por excelência, o morfema. Com isso o morfema passava a ter lugar no *léxico*: “O estoque total de morfemas numa língua é seu léxico” (Bloomfield, 1933: 162).

São os morfemas os elementos centrais para o modelo; a *palavra*, do mesmo modo que o sintagma, resulta dessas combinações. Tais combinações são arranjos hierárquicos de constituintes imediatos, como exemplificado em Hockett (1958: 180¹² e 154¹³), reproduzidos em (3.7a, b) respectivamente:

(3.7) a.



b.



Uma vez que a unidade básica da **gramática** (isto é, da morfologia e da sintaxe) é o morfema, a distinção entre morfologia e sintaxe não é nítida, como se pode cons-

¹¹ Bloomfield (1933: 163): “Os arranjos significativos de formas numa língua constituem sua *gramática*”. E ainda (*id.* 163ss): “As formas lingüísticas podem apresentar quatro tipos de arranjos: *ordem, modulação, modificação, fonética, seleção*”.

¹² Traduzível como ‘descortesia’.

¹³ Traduzível como ‘A Inglaterra utiliza o sistema pé-libra’.

tatar em (3.7): os constituintes do primeiro nível do diagrama tanto podem ser palavras como os elementos que as constituem. O próprio Hockett (1958: 178) “ajustaria” a definição de **morfologia** para “todas as construções em que um CI é menor do que uma palavra...”. Ou, na pena de Gleason Jr. (1961: 137), trecho já referido anteriormente (seção 2.3), “Nem sempre é clara a distinção entre morfologia e sintaxe. [...] Não é, porém, possível uma discriminação mais satisfatória que abranja as línguas em geral”. Reconhece-se, no entanto, que as construções morfológicas são mais elaboradas que as sintáticas (Bloomfield, 1933: 207).

3.3.1. Os tipos de morfemas

Até aqui os exemplos que apresentamos constituem-se de um *radical* e *afixos*, que se combinam na formação de uma palavra. Uma das conseqüências de se trabalhar com um modelo baseado na noção de *item* (e não na de *palavra*) foi a de que, de algum modo, se passou a compreender a morfologia sintagmaticamente, como *afixação*.

A análise de diferentes línguas demonstrou, porém, que os processos morfológicos podem ser de outros tipos, captados não tão elegantemente nessa análise de um significado para um significante. Mecanismos como *morfemas alternativos, subtrativos* e *zero*, embora considerados “uma difícil manobra, contudo necessária” (Hockett, 1947: 230), vieram somar-se, assim, ao *morfema aditivo*. Apresentamos em seguida os tipos de morfemas.

3.3.1.1. Morfema aditivo

Os **morfemas aditivos** são facilmente captados pela análise morfêmica: são os *radicais* e os *afixos*. Os exemplos que vimos até aqui (*am-*, *-s*) incluem-se neste grupo.

A **raiz** ou **radical primário** é o elemento mínimo de significado lexical. Se for ampliado por derivação ou por composição, forma o **radical** ou **radical secundário**. Em *transformar*, por exemplo, temos um radical ou radical secundário *transform-*. Retirado o prefixo (*trans-*), ficamos com a raiz ou radical primário *form-*. Em *terceiro-mundista*, temos um radical, (*terceiro-mundo*) que é a base do derivado em *-ista*; temos, porém, duas raízes.

No estudo da formação de palavras encontra-se muitas vezes o termo **base** em lugar de *raiz* ou de *radical* ou ainda de *palavra*. Retomando os exemplos acima, *transform-* e *terceiro-mundo* são as bases, respectivamente, de *transformar* e *terceiro-mundista*.

O que chamamos **afixos** são elementos que se distinguem pela posição que tomam em relação à raiz. Os **prefixos** antecedem a raiz, como *re-* em *reler*; os **sufixos** seguem a raiz, como *-al* em *arrozal*.

A morfologia das línguas costuma repousar predominantemente em apenas um desses dois tipos de afixos. Afinal, “uma língua com dez classes de posições tanto para prefixos como para sufixos traria dificuldades reais para a localização das raízes das palavras” (Anderson, 1985b: 166-167). Há tendência entre as línguas para maior número de sufixos: “Línguas exclusivamente sufixais são razoavelmente comuns, ao passo que línguas exclusivamente prefixais são bastante raras” (Greenberg, 1963:92).

Uma das generalizações acerca das línguas do mundo diz respeito à ordenação que sufixos (ou prefixos) derivacionais e sufixos (ou prefixos) flexionais tomam em relação à raiz. Essa generalização foi expressa por um dos *universais implicacionais* de Greenberg (1963: 93):

Universal 28. Se tanto a derivação como a flexão seguem a raiz, ou ambas precedem a raiz, a derivação está sempre entre a raiz e a flexão.

Vamos a um exemplo: *livr-inho-s* apresenta mais perto do radical *livr-* o sufixo derivacional *-inho* que a marca de Plural.

☞ Não são tão claros casos como os dos plurais de alguns diminutivos portugueses, tais como *pãezinhos*, *florezinhas*, *pasteizinhos*. Ou ainda dos derivados em *-mente* cujas bases são adjetivos variáveis quanto ao Gênero: *lindamente*, *historicamente*...

Os **infixos** são um tanto diferentes dos *prefixos* e dos *sufixos*, uma vez que se inserem por completo no interior de uma raiz, tornando-a *descontínua*. Em árabe, como de resto nas línguas semíticas, o padrão mais geral para as raízes é CCC (em que C equivale a *consoante*), que recebe infixos (dados em Nida, 1949: 68, que remete a Bloomfield, 1933):

(3.8) *ktb	‘escrever’
katab	‘ele escreveu’
katib	‘escrevendo’
kitab	‘livro’
maktab	‘lugar para escrever’

Note-se que a característica definidora do infixo é a de ele posicionar-se no interior de uma raiz. Não se entenda, portanto, como infixo um afixo que se apresente entre outros dois afixos. Vamos a um exemplo. Em turco, o verbo *yemek* ‘comer’ apresenta, no seu imenso paradigma de cerca de três mil formas, a forma *yedim* ‘comi’ e outra, *yemedim* ‘não comi’. Não há razão para se considerar *-me-* um infixo. É um sufixo numa seqüência de sufixos, como ilustrado a seguir (*vide* Gleason, 1961: 121):

(3.9) ye-	me-	di-	m
comer	NEG	PAS	1SG/AG

☞ Não são apenas as raízes que são focalizadas como podendo apresentar-se como formas descontínuas, em virtude de um morfema ter sido incluído no seu interior. Segundo Nida (1949: 67-68; 76), em zoque, língua indígena do México, por exemplo, vários morfemas apresentariam alomorfes descontínuos, numa análise bem ao gosto de IA. Comparem-se as formas básicas em (a) com as seqüências em (b):

(a) ken-	‘olhar’	(b) kenu	‘olhar-PAS’
y-	‘3SG/ S/ TRANS’	kenpa	‘ele olha, ele olhará’
-hay-	‘BEN’ ¹⁴	kyenhayu	‘ele viu isso para ele (em seu favor)’
-to ³ y-	‘DES’ ¹⁵	ken ³ o ³ yu	‘ele queria olhar’
-u	‘PAS’	ken ³ o ³ pya	‘ele quer olhar’
-pa	‘PRES OU FUT’		

O Benefactivo e o Presente/Futuro podem, por exemplo, apresentar-se na seqüência *-hapy a-*, em que cada um está parcialmente dentro do outro.

Também os morfemas alternativos fazem um morfema apresentar-se numa forma descontínua. Como dissemos anteriormente, casos como esses não se ajustam bem à análise em constituintes imediatos, embora evitem que se fale em *regras* ou *processos*.

por meio de um sufixo e a adjução de um prefixo, este sem significação própria. Uma alternativa de análise para que se evite a postulação de morfemas sem significado e para levar em consideração a adição simultânea de elementos nas posições inicial e final da formação – que é a característica definidora dos derivados parassintéticos – é considerar formas como *amanhecer* com a estrutura em (3.10):

(3.10) manhã + a....ec(e(r)).

Neste caso, *a....ec(e(r))* é um afixo descontínuo, o **circunfixo** (e por essa razão a denominação *circunfixação* é empregada em lugar de *parassíntese* caso se leve em conta tal proposta).

Para alguns estruturalistas, como Nida, por exemplo, os *morfemas reduplicativos*, de que trataremos a seguir, incluem-se entre os morfemas aditivos.

3.3.1.2. Morfema reduplicativo

O **morfema reduplicativo** pode ser focalizado não como um afixo com formas bastante variáveis, mas como uma modificação na raiz, que consiste na repetição de toda ela ou de parte dela.

¹⁴ O Benefactivo é o Caso que indica em favor de quem algo é feito.

¹⁵ Desiderativo é o Modo para a expressão do desejo ou vontade.

De acordo com Anderson (1985b: 169), são quatro os tipos mais comuns de reduplicação. A cópia pode afetar a consoante ou grupo inicial. Nesses casos não é difícil encontrar-se a inserção de uma vogal determinada. Alguns perfeitos latinos, por exemplo, são marcados pela repetição da consoante inicial do primeiro radical do verbo, seguida de *-e-*:

(3.11) Presente		Perfeito
pango 'concordo'		pepĩgĩ
pargo 'abstenho-me de'		peperci
cano 'canto, celebro'		cecĩnĩ

A cópia pode afetar a primeira seqüência inicial de consoante e vogal: no pídgín da Nova Guiné, *pairap* 'explodir' é intensificado em *papairap* 'explodir diversas vezes e em rápida seqüência' (Mühlhäusler, 1979: 405). Pode também copiar toda a primeira sílaba, ou ainda toda a raiz, como no pídgín da Nova Guiné *pretpret* 'estar muito amendrontado', formado a partir de *pret* 'estar amendrontado' (Mühlhäusler, 1979: 406).

A reduplicação não afeta necessariamente apenas seqüências fônicas iniciais. Ainda no pídgín da Nova Guiné é possível encontrar a reduplicação da sílaba final, por exemplo: *lapun* 'velho', mas *lapunpun* 'muito velho'. Observem-se também os exemplos a seguir, acerca do Número no verbo em samoano, língua austronésia falada em Samoa (oceano Pacífico – Gleason, 1955: 29; Jensen, 1990: 70):

(3.11)	manao ¹⁶	'quer'	mananao	'querem'
	matua	'é velho'	matutua	'são velhos'
	malosi	'é forte'	malolosi	'são fortes'
	punou	'retesa'	punonou	'retesam'
	pese	'canta'	pepese	'cantam'
	alofa	'ama'	alolofa	'amam'
	galue	'trabalha'	galulue	'trabalham'
	maliu	'morre'	maliliu	'morrem'

No exemplo acima, a expressão do Número Plural se faz pela repetição da sílaba mais próxima àquela que é a última na forma singular. Jensen (1990: 71) analisa esses exemplos como um *processo*, formalizado como em (3.13), a seguir, em que o símbolo X indica que não importa o que veio antes da penúltima sílaba, isto é, da penúltima seqüência *C(onsoante)+V(ogal)*; o colchete de fechamento J indica o final do vocábulo; o _v em letra subscrita indica que a classe gramatical da palavra é *verbo*:

(3.13)	X	C	V	(C)	V] _v										
	1	2	3	4	5	6	→	1	2	3	2	3	4	5	6
															[+ plural]

¹⁶ As seqüências de vogais contam como sílabas separadas. Não há ditongos nessa língua (Jensen, 1990: 71).

Observe como o tratamento que Gleason Jr. (1961: 96) apresenta para a reduplicação é diferente daquele formalizado por Jensen e apresentado em (3.13). Jensen, ao empregar uma regra para descrever a formação do Plural, capta uma generalização acerca da gramática dessa língua. Gleason Jr., por seu turno, define a **reduplicação** como "afixos com formas extremamente variáveis". Estamos, assim, na análise de Gleason Jr., perante uma lista de morfemas aditivos. Ao apresentar a reduplicação não como uma regra, mas como uma lista de afixos, Gleason Jr. mantém-se atrelado a IA, mas perde em generalização. Vejamos um dos exemplos que apresenta, do tagalo, língua austronésia falada nas Filipinas (Gleason Jr. *id. et ibid*):

(3.14)	/isá/	'um'	/iisá/	'só um'	prefixo /i-/
	/dalawá/	'dois'	/dadalawá/	'só dois'	/da-/
	/tatló/	'três'	/tatatló/	'só três'	/ta-/
	/píso/	'peso'	/pipíso/	'só um peso'	/pi-/ etc.

Diferentemente de Jensen, que muda uma forma, básica, em outra, Gleason Jr. apresenta-nos uma lista de alomorfes (a parte que se repete) que expressam a mesma noção, embora com formas fonológicas bem distintas.

Anderson (1985b: 170) observa que é razoavelmente limitado o conjunto de significados expressos pela reduplicação. Nos nomes indica "pluralidade, formas diminutivas (ou aumentativas), mas não caso, gênero ou categorias dêiticas/referenciais". Nos verbos, a reduplicação costuma indicar diferenças aspectuais ("tais como progressivos, imperfeitos, perfeitos representando um estado, distributivos, interativos etc.), formas plurais e modos tais como hipotético, não realizado etc., mas não pessoa, voz...", ou mesmo tempo, se esta categoria está dissociada de oposições aspectuais.

3.3.1.3. Morfema alternativo

O **morfema alternativo** consiste na mudança da estrutura fônica da raiz (seja por alternância da qualidade ou quantidade de vogais, de consoantes, de acento ou de tom), como em *digo/dizes*. Exemplos clássicos desses morfemas nas línguas ditas flexivas são o *Ablaut* ou *apofonia* e o *Umlaut* ou *metafonia*, dois processos de alternância de vogais no interior de uma raiz, cuja distinção se faz, em geral, na lingüística histórica. Em grego clássico, por exemplo, a alternância entre /e/, marca o presente, e /o/, aliada à reduplicação, o perfeito: *leipo/leloipa*; em português, casos como *fiz/fez*.

Muito frequentemente as alterações na raiz são acompanhadas por outros afixos. É, por exemplo, o padrão geral do Plural de nomes em inglês que explica a postulação de um alomorfe zero de Plural, além da alternância, para a palavra inglesa *feet* 'pés' (plural de *foot* 'pé'). Segundo Nida (1949: 54), *feet* é decomponível em três morfemas: (a) o radical; (b) a substituição de /u/ por /iy/; e (c) o sufixo zero, este último, na realidade, um alomorfe zero de Plural (Nida, 1949: 46n44). À substituição de /u/ por /iy/ é atribuído o *status* de morfema porque se constitui na única diferença visível entre o Singular e o Plural (Nida, 1949: 54).

1 Vaca é feminino de boi?

A literatura lingüística refere exemplos como *boi/vaca, sou/fui/era* sob a denominação **supletivismo**, que é a alteração máxima que se poderia aplicar a uma raiz: substituí-la por outra forma. Uma alternativa a esse enfoque flexional é considerar que estamos diante de itens lexicais diferentes, relacionados semanticamente, cada um expressando diferentes propriedades.

O motivo que levou à inclusão desses casos entre os fenômenos flexionais nos estudos gramaticais foi, justamente, o estudo de línguas morfologicamente muito flexionadas como o latim, o grego e o sânscrito. Na medida em que lidava com línguas que apresentavam processos de alteração das raízes, a gramática tradicional focalizou tais exemplos como os casos mais radicais desses processos.

Um exemplo clássico pode tornar mais claro o nível de alteração das raízes estudado nas gramáticas tradicionais: o sistema de **Ablaut** (ou alteração de vogais) do proto-indo-europeu, aqui exemplificado com a raiz indo-européia **bher* 'carregar' (Hock, 1986: 545). As alterações são classificadas em graus:

- grau normal, ou grau-e**, em que se mantém o *-e-* na raiz, como em **bher-ō*: gr. *pher-ō*; lat. *fer-ō* 'eu carrego';
- grau-o**, em que há a substituição do *e* do grau normal por *o*, como em **bhor-eyō*: gr. *phor-eō* 'carrego repetidamente';
- grau expandido**, com a substituição de *e* ou *o* pela vogal longa correspondente, como em **e-bher-s-t*: sânscr. *a-bhā-r* 'carregou'; gr. *phōr* 'ladrão';
- grau -φ**, em que há a supressão da vogal básica, como em **bhr-ti*: sânscr. *bhr-ti* - 'um carregamento'.

3.3.1.4. Morfema zero

O morfema zero é de natureza diferente. Observem-se os exemplos em (3.15), referentes ao Gênero do nome (mais especificamente, de nomes relativos a seres animados) em português.

(3.15) Masculino	Feminino
mestre	mestra
hóspede	hóspeda
leitor	leitora
professor	professora
marquês	marquesa
menino	menina
gato	gata

A marca de Feminino é sempre */-a/* nesses exemplos. Sua ausência é significativa como característica de Masculino. Daí Camara Jr. ter postulado um **morfema zero** para o Masculino em português, isto é, "um morfema no qual não haja nenhum alomorfe evidente" (Gleason Jr., 1961: 80).

A postulação do morfema zero não foi aceita por todos os lingüistas. Gleason Jr. (*id. et ibid.*), por exemplo, considerou-o "absolutamente desnecessário" e "logicamente insustentável", porque não haveria "fronteiras definíveis" para a justificativa de seu emprego, uma vez que poderíamos "adicionar livremente zeros de toda a ordem às nossas descrições, cada um deles tão defensável como o último". Uma forma como *mestre*, por exemplo, poderia ser analisada como (3.16):

(3.16) mestre + \emptyset + \emptyset
 MASC SG

A crítica de Gleason Jr. torna-se mais aguda se atentarmos para a diferença entre fenômenos flexionais e derivacionais. As categorias gramaticais formam classes fechadas. Ao trabalhar com elementos flexionais, o lingüista postula zeros que se opõem a elementos bem definidos, representantes das propriedades morfossintáticas. Ao lidar com dados derivacionais, porém, apresenta-se-lhe quadro diverso. Poderia opor a marca de Masculino à de Feminino, por exemplo; mas que poderia opor, digamos, a uma derivação em *-dade*? Ou melhor: o fato de uma forma não apresentar um derivado em *-dade* seria motivo para a postulação de um morfema zero? Os dados não permitiriam, neste caso, um uso justificável desse artifício descritivo.

Mesmo aqueles que não se opuseram à postulação de morfemas zero advertiram quanto à parcimônia com que deveriam ser empregados. É o caso de Nida (1949: 46), por exemplo: "Deve-se evitar [...] o uso indiscriminado de zeros morfêmicos. De outro modo a descrição de uma língua torna-se excessivamente polvilhada de zeros, devidos apenas à congruência estrutural e ao equilíbrio".

3.3.1.5. Morfema subtrativo

O **morfema subtrativo** consiste num radical que perde fonemas para a expressão de um dado traço gramatical (Nida, 1949: 75). O exemplo clássico do morfema subtrativo foi apresentado por Bloomfield (1933: 217) para o Gênero do adjetivo em francês:

(3.17) Masculino			Feminino	
plat	[pla]	'plano'	platte	[plat]
laid	[le]	'feio'	laide	[lɛd]
distinct	[diste ⁿ]	'distinto'	distincte	[diste ⁿ kt]
long	[lo ⁿ]	'longo'	longue	[lo ⁿ g]
bas	[ba]	'baixo'	basse	[ba:s]
gris	[gri]	'cinza'	grise	[gri:z]
frais	[frɛ]	'fresco'	fraîche	[frɛ:ʃ]
gentil	[žɑ ⁿ ti]	'gentil'	gentille	[žɑ ⁿ ti:j]
léger	[leže]	'luz'	légère	[leže:r]
soul	[su]	'bêbado'	soule	[sul]
plein	[ple ⁿ]	'cheio'	pleine	[ple:n]

Uma possibilidade de análise seria a de se tratar o Feminino como um caso de morfema aditivo: a uma forma básica masculina somar-se-ia uma consoante, indicadora de Feminino. O problema seria: que consoante? As consoantes finais da forma feminina somente são previsíveis caso se tome a escrita como base (mas este é um procedimento inválido, como vimos anteriormente). Daí a proposta de Bloomfield (*id. et ibid.*):

se tomamos a forma feminina como nossa base, podemos descrever esse tipo irregular pela assertiva simples de que a forma masculina é derivada da feminina por meio de um traço subtrativo (ing. minus-feature), a saber, a perda da consoante final ou do grupo [-kt].

O morfema subtrativo, juntamente com o alternativo e o zero, faz parte do grupo dos “malcomportados” (Anderson, 1988: 153). Todos ajustam-se mal a IA, porque não há como segmentar “a forma de superfície de uma palavra contendo essa categoria de modo que alguma subparte (possivelmente descontínua) de sua estrutura constitua o morfe em questão” (Anderson, 1988: 160). Mas o que é um morfe?

3.3.2. O morfema é uma classe de morfes

O morfe é um segmento de enunciado, ou melhor, uma seqüência fônica, a que é possível atribuir significado e que será posteriormente classificado num morfema. O morfema é, por conseguinte, uma abstração em relação ao morfe, do mesmo modo que o fonema o é em relação ao fone: um morfema é uma classe de morfes, isto é, cada morfe, ou alternante morfêmica, é um elemento de um conjunto (que pode ser unitário) formador de uma unidade estrutural, que é o morfema. Qualquer enunciado é completamente composto de morfes (Hockett, 1947:230). Também são morfes todas as seqüências fonêmicas que restam após a divisão do enunciado, tenham ou não significado (Hockett, 1947: 239).

Quando, anteriormente, segmentamos *gatinhos*, segmentamos esse enunciado em morfes, não em morfemas. Qual a diferença? Tomemos para exemplo o -s final, que identificamos como marcador de Plural. Se ampliássemos nosso corpus, de modo que, além dos exemplos acima, ele incluísse também formas como as de (3.18), a seguir, veríamos que o Plural em português pode manifestar-se ainda de outras maneiras que não a adição de /-s/. Vejamos algumas delas:

(3.18) Singular	Plural	Singular	Plural
a. cruz	cruzes	b. lápis	lápiz \emptyset
pilar	pilares	cactus	cactus \emptyset
líquen	líquenes	pires	pires \emptyset

Os poucos exemplos acima revelam que /-s/ não é a única expressão possível para o Plural em português. Em outras palavras: o morfema de Plural realiza-se (pelo

menos) por meio de dois outros *alomorfes*, cada um deles com uma distribuição¹⁷ determinada. Assim, acrescenta-se /-s/ ao nome no Singular se termina em vogal ou ditongo; /-es/, caso termine em consoante, como em (3.18a) acima, exceção feita a nomes paroxítonos em /s/, quando então a forma permanece invariável, como em (3.18b). Neste último caso estaríamos perante um **alomorfe zero**: dentre as possíveis realizações de um morfema uma delas é não apresentar realização fonêmica. O morfema de Plural não é, portanto, a realização específica x ou y, mas o conjunto dessas realizações, ou *alomorfes*. Poderíamos representá-lo como em (3.19), com o elemento que aparece no maior número de ambientes tomado para a representação do morfema. As reticências indicam que, com a ampliação do corpus, poderíamos encontrar outros alomorfes:

$$(3.19) \{S\} \rightarrow /-s \sim -es \sim \emptyset \sim \dots/$$

PL

3.3.3. A alomorfia

Como dissemos anteriormente, para chegar aos morfemas de uma língua o linguísta procede pela substituição, já adotada na fonêmica: isola-se o *foco* da análise do restante do ambiente, como nos experimentos controlados das ciências naturais (Haugen, 1951:360). Essa técnica permite ao pesquisador segmentar o enunciado em morfes, que serão posteriormente classificados em morfemas, a partir de procedimentos bem determinados, como, por exemplo, os princípios estabelecidos por Nida (1949), que reproduzimos adiante, no quadro 5.

Cabe ao linguísta comparar enunciados parcialmente diferentes, para formas que apresentem distintividade semântica comum, em busca de identidade fonêmica (*princípio a*) ou de diferenças que possam ser definidas fonemicamente (*princípio b*), de distribuição complementar ou variação livre (*princípio c*).

Como identificar, por exemplo, os elementos formadores do vocábulo *gatinhos*? A resposta é: com a ajuda de vocábulos parcialmente semelhantes. Pela comparação de *gatinhos* com *gatinho* e com pares como *pato/patos*, *mestre/mestres*, depreende-se o -s final como marca de Plural. Pela comparação com *gato* e com formas como *patinho*, *ursinho* isola-se mais um elemento: *-inho*, indicador de diminutivo. Resta a forma *gat-*, que é recorrente e que se não consegue, por sua vez, subdividir em unidades menores com significado: compare-se *gatinho* com formas parcialmente semelhantes como *gato*, *gataria*.

E *ágata*, por exemplo? Não é possível comutar *ágata* com *gato* sem que se destrua qualquer relação de significado. Uma vez que aquela forma ocorre em isolado, é classificável num morfema (*princípio f*).

Deve-se ter sempre em mente que a seqüência sonora tem de estar associada a um determinado significado. É este um dos fundamentos da análise: o morfema é

¹⁷ Por **distribuição** entenda-se o conjunto de ambientes em que uma forma pode ocorrer.

uma “unidade mínima com significado” (Nida, 1949: 6). Ora, se a mesma seqüência de fonemas /gat/ está presente em termos como *ágata*, *gaturamo*¹⁸, isto não é condição suficiente para classificarmos tais ocorrências em conjunto com a seqüência /gat/ de *gato*. É que não se pode aí identificar qualquer semelhança de significado com a forma {GAT-} presente em *gato*, *gatinho*, *gatão* etc., e, por conseguinte, subdividir as formas *ágata* e *gaturamo* (em *á-gat-a* ou *gat-uramo*), comutando-as com *gato*.

E formas como *canto* ‘atividade de cantar’ e *canto* ‘junção de duas paredes’? Novamente, embora fonemicamente idênticas, tais formas não podem ser relacionadas, porque seus significados nada têm em comum. Trata-se de formas homófonas, que não serão classificadas no mesmo morfema (*princípio e*).

Os estruturalistas tiveram dificuldade em lidar com a noção de significado. Joos (1950: 356), por exemplo, define o significado de um morfema em termos distribucionais: o significado de um morfema é, “por definição, o conjunto de probabilidades condicionais de sua ocorrência em relação a todos os outros morfemas”. Para alguns problemas que cercaram o tratamento estruturalista do significado, a partir da noção de *raiz*, ver *Operacionalização do conceito de raiz* (Basilio, 1974b).

Quadro 5: Seis princípios para a identificação dos morfemas de uma língua (Nida, 1949).

- a) “Constituem um morfema único as formas que têm distintividade semântica comum e forma fonêmica idêntica em todas as suas ocorrências”;
- b) “As formas que apresentam distintividade semântica comum mas diferentes na forma fonêmica (*i.e.*, os fonemas ou a ordem dos fonemas) podem constituir um morfema desde que a distribuição de diferenças formais seja fonologicamente definível”;
- c) “As formas que apresentam distintividade semântica comum mas que diferem na forma fonêmica de tal modo que sua distribuição não pode ser fonologicamente definida constituem um morfema único se estão em distribuição complementar em acordo com as seguintes restrições:
 1. A ocorrência nas mesmas séries estruturais têm precedência sobre a ocorrência em diferentes séries estruturais na determinação do *status* de morfema;

¹⁸ Observe-se que, conforme já vimos anteriormente (cap. 1), não há a possibilidade de se argumentar que *gat-* em *gato* e em *gaturamo* não podem ser classificados num mesmo elemento por terem étimos diferentes: respectivamente, o latim *cattus* e o tupi *catarama*.

2. A distribuição complementar em séries estruturais diferentes constitui a base para a combinação de possíveis alomorfes num morfema apenas se aí também ocorrer nessas séries estruturais diferentes um morfema que pertença à mesma classe de distribuição, como as séries alomórficas em questão, e que ele próprio tenha apenas um alomorfe ou alomorfes definidos fonologicamente.
 3. Ambientes táticos imediatos (ing. *immediate tactical environments*) têm precedência sobre ambientes táticos não-imediatos na determinação do *status* morfêmico.
 4. Contraste em ambiente distribucional idêntico pode ser tratado como submorfêmico se a diferença no significado dos alomorfes reflete a distribuição dessas formas.
- d) “Uma diferença formal explícita numa série estrutural constitui um morfema se em qualquer membro dessa série a diferença formal explícita e uma diferença estrutural zero são os únicos traços para distinguir uma unidade mínima de distintividade fonético-semântica.” (*v. seção 3.3.1.3*);
- e) “As formas homófonas são identificáveis como o mesmo morfema ou como morfemas diferentes com base nas seguintes condições:
1. As formas homófonas com significados nitidamente diferentes constituem morfemas diferentes.
 2. As formas homófonas com significados relacionados constituem um único morfema se as classes de significado podem ser postas em paralelo pelas diferenças distribucionais, mas constituem múltiplos morfemas se as classes de significado não podem ser postas em paralelo pelas diferenças distribucionais.”;
- f) “Um morfema é isolável se ocorre sob as seguintes condições:
1. Em isolado.
 2. Em combinações múltiplas, pelo menos em uma das quais a unidade com que se combina ocorre em isolado ou em outras combinações.
 3. Numa combinação única, desde que o elemento com o qual se combina ocorra em isolado ou em outras combinações com constituintes não-únicos.”

3.3.4. A morfotática

Se a **alomorfia** diz respeito às configurações que um morfema pode tomar, a **morfotática** dá conta das restrições à combinação de morfemas, ou melhor, estuda a sua distribuição. A difusão do termo deve-se a Hockett: por **tática** (< ingl. *tactics*) compreende-se qualquer teoria sobre a combinação de unidades similares, como fonemas com fonemas¹⁹, morfemas com morfemas.

A morfotática leva em conta: *formas, ordem (linear), construções e hierarquia de constituintes*, representada por CI.

¹⁹ Seria a fonotática.

- 1 Segundo Hockett (1954: 389), o *padrão tático* de uma língua resulta de:
- Uma lista das construções.
 - Em cada construção:
 - a enumeração das posições nela contidas;
 - a especificação de algum marcador para essa construção;
- Para cada posição:
- uma lista dos morfemas que nela podem ocorrer;
 - uma lista das construções que nela podem ocorrer.

Para estabelecer o padrão tático de uma língua no tocante à morfologia, descrevem-se as classes de distribuição característica dos morfemas. As "classes mais universais em todas as línguas do mundo são a dos *radicais* e a dos *afixos*" (Gleason Jr., 1961: 63). Os radicais e os afixos distribuem-se, por sua vez, em subclasses. Dentre os radicais, há, por exemplo, radicais verbais e radicais nominais; dentre os afixos, encontramos, por exemplo, os sufixos. Um sufixo segue um radical, o que faz com que **docaminha*, por exemplo, seja impossível para a expressão do participio de CAMINHAR em português.

3.4. A morfofonêmica

Alguns estruturalistas norte-americanos, como Hockett e Wells, por exemplo, defenderam a necessidade de um nível intermediário entre a morfologia e a fonologia. A **morfofonêmica** é "o ramo da gramática que lida com a forma fonêmica de morfemas, palavras e construções, sem atenção ao significado" (Hockett, 1942: 107). Esse nível segue-se da inclusão de uma **condição de biunivocidade** na análise fonêmica. Tal condição previa que a uma determinada cadeia de fonemas deveria corresponder uma cadeia de fones e vice-versa. Assim, para uma forma como *asa* teríamos a correspondência fonema/fone como ilustrado a seguir:

(3.20)	/a/	+	/z/	+	/a/	seqüência fonêmica
	↕		↕		↕	
	[a]		[z]		[ə]	seqüência fonética

Podemos afirmar que /z/ é um fonema no português porque pode opor formas na língua (*asa/assa*, por exemplo). Pela mesma razão, também /s/ pode ser considerado um fonema: *acha/assa/asa*. Pela condição da biunivocidade, dada uma cadeia fonética, deveria também ser possível depreender a cadeia fonêmica. Imagine, então, a seqüência fonética *casas*, mas seguida de uma palavra começada por vogal, como em *casas alegres*. Deveria corresponder a algo como (3.21):

(3.21)	[k]	+	[a]	+	[z]	+	[ə]	+	[z]	+	[a]	+	[l]	+	[ε]	+	[g]	+	[r]	+	[i]	+	[š]	
	↕		↕		↕		↕		↕		↕		↕		↕		↕		↕		↕		↕	
	/k/		/a/		/z/		/ə/		/z/		/a/		/l/		/ε/		/g/		/r/		/e/		/š/	

?

Eis aqui um problema, apontado pelas setas. Havíamos afirmado que tanto /z/ como /š/ eram fonemas, bem como /s/, e *uma vez fonema, sempre fonema*²⁰. No entanto, nessa posição, tais elementos não contrastam: podemos ter *casa[z]* (*alegres*), *casa[š]* (*verdes*), *casa[š]* (*feias*). A saída para aqueles que adotaram esse enfoque foi a postulação de um terceiro nível de análise, intermediário. Listam-se os morfofonemas (representados entre barras verticais, assim | |), os fonemas para substituí-los e os ambientes em que ocorrem. Pontes (1965: 32), por exemplo, ao tratar dos verbos, propõe para o -s final, que ela representa pelo morfofonema |z|²¹, a seguinte lista de fonemas: (a) /z/ diante de vogal; (b) /š/ diante de silêncio ou consoante surda; e (c) /ž/ diante de consoante sonora. A inclusão de mais um nível de análise foi o preço por procurar levar IA às últimas conseqüências.

Vimos até aqui o modo como uma influente corrente da lingüística compreendeu os princípios que regem a estrutura da palavra. Passamos, em seguida, a alguns dos problemas com esse tipo de análise.

3.5. Alguns problemas para a análise morfêmica

Considerar o morfema como a unidade mínima de som e significado em que se baseia a análise acarretou problemas, que não passaram despercebidos já aos lingüistas da época. Apresentamos, em seguida, quatro desses problemas, todos recorrentes na literatura: os *fonestemas*, os *morfes supérfluos*, os *morfes vazios*, os *morfes cumulativos*.

²⁰ Não se pode, ao adotar a biunivocidade, lançar mão do conceito de arquifonema, possível em IP.

²¹ Pontes (1965: 31): "É claro que os morfofonemas estabelecidos, foram-no com base no paradigma verbal: correspondem a alternâncias fonologicamente condicionadas de morfemas verbais. Foi pela observação dos paradigmas que os estabelecemos. É possível que algum não se aplique a outras classes de palavras." A escolha de |z| é explicada em nota (*id.*:46n3): "...diante de vogal poderia ocorrer qualquer um dos fonemas, mas, diante de consoante, /z/ não ocorre (no português do RJ): sendo ela surda, só encontramos /š/ e sendo sonora, somente /ž/. A troca é automática, portanto".

3.5.1. Os fonestemas ou elementos fonestéticos ou simbolismo fonético

Os **fonestemas** (< ingl. *phonaesthemes* < *phono-* 'som' + *aesthé-* 'perceber' + *-ema* 'unidade estrutural') são seqüências sonoras recorrentes que associam algumas poucas palavras cujo significado é relacionado (Joseph, 1998: 360). O exemplo clássico é a seqüência *gl* em vocábulos ingleses semanticamente relacionados como *glimmer* 'reflexo', *gleam* 'brilho', *glow* 'fulgor', *glitter* 'cintilação'. Pode-se apontar como um exemplo paralelo em português a nasal /n/ em formas como *nada*, *ninguém*, *nenhum*, *não*, *nenhures*. Em geral tais seqüências não foram consideradas morfêmicas, embora fossem recorrentes e se lhes pudesse atribuir algum tipo de identidade de significado. O ponto em questão é o de que, ao se reconhecerem tais formas no interior de unidades maiores, destroem-se, para o restante do vocábulo, as relações de som e significado que fundamentam a segmentação em unidades mínimas.

① Nida argumenta que tais formas não podem ser isoladas como morfemas "uma vez que não ocorrem como formas livres ou com formas que ocorram em outras combinações". Já Joseph (1998:361) defende algum *status* gramatical para os fonestemas com base na diacronia do inglês. A seqüência *-ag* aparece em palavras inglesas cujo sentido poderia ser glosado como 'ação tediosa, cansativa ou lenta', como em *drag* 'arrastar (com força e dificuldade)', *flag* 'fatigar, estafar', *flag* 'esmorecer', *lag* 'demorar-se', todas elas já presentes no inglês médio (ca. 1100-1500). Essas formas, segundo o A., teriam atraído para sua órbita *sak*, que, a partir do século XVI torna-se *sag* 'afundar, descair'.

3.5.2. Os morfes supérfluos (Anderson, 1992:54)

É possível em português formar advérbios a partir de adjetivos, juntando a estes o sufixo *-mente*: *doce/docemente*, *triste/tristemente*, *feliz/felizmente*. Os adjetivos *doce*, *triste* e *feliz* são uniformes: *homem doce/triste/feliz*, *mulher doce/triste/feliz*. Mas e quando não o são? Bem, nesses casos os advérbios em *-mente* têm de ser derivados a partir da forma de Feminino do adjetivo: *lindamente*, *graciosamente*, *fonologicamente*. Existe uma razão histórica para isso: *-mente* deriva do substantivo feminino *mente*, com que o adjetivo que o antecedia concordava.

Deixando de lado a história da língua, que não pode, numa análise sincrônica, ser chamada à cena, temos três morfemas em *lindamente*: *lind-* + *-a* + *-mente*. Que fazer com *-a* nesse tipo de análise? Podemos classificá-lo como marca de Feminino, mas tal marca é decididamente supérflua num vocábulo invariável e, além disso, incompatível com o significado do vocábulo. Se a considerarmos como expressão do Gênero, temos ainda um problema a mais, que é o de marcar uma flexão (supérflua) antes da derivação. Resta-nos ainda uma outra alternativa: classificar esse *-a-* como um *morfe vazio*: *-a-* não seria considerado marca de Feminino, mas algo como uma vogal de ligação, sem significado, por conseguinte, e teríamos tentado sair de um problema criando outro, como veremos em seguida.

3.5.3. O morfe vazio

Um dos problemas para a análise estruturalista são os elementos recorrentes que não apresentam significado. Um **morfe vazio** é um morfe que não tem significado e que não é atribuído a nenhum morfema. Para Hockett (1947: 238), /a/, vogal temática do infinitivo espanhol *amar* é um morfe vazio, uma vez que a forma pode ser analisada como composta da raiz *am-* mais a terminação *-r* de infinitivo²².

Se "o morfe vazio não é atribuído a morfema algum, não tem razão de ser" (Basílio, 1974a: 83). Que diferença pode ser postulada entre esse /a/ sem significado e que não é atribuído a morfema algum, por exemplo, e o fonema /a/? Caso aceitemos a proposta de morfemas vazios, a definição do morfema como elemento mínimo de som e significado desmorona. É como um cobertor curto: ou ficam dados a descoberto, ou fica-se com princípios de análise que são inúteis, uma vez que só funcionam às vezes.

3.5.4. O morfe cumulativo

Na análise morfêmica espera-se que a um elemento de significado deva corresponder um elemento no nível da expressão e vice-versa. Retomemos, no entanto, a análise de Pontes (1965) para a forma portuguesa *amas*, apresentada antes, em (3.5b) e aqui repetida por conveniência:

(3.5) b.

amas		
am-	-a-	-s
raiz	VT + IND. PRES.	2SG

Ao propor que um único morfe pudesse representar a vogal temática e a desinência para TMA (as quais, em outras formas do verbo, tais como em *am-á-va-mos*, são elementos separados) quebrava-se esse desiderato: um único morfe representava duas posições distintas do padrão verbal. O mesmo em sua análise para *amo* [que segue a de Hockett (1947)]: |-o| indica, cumulativamente, IND. PRES. + 1SG.

Problema semelhante apresenta-se no caso da fusão de dois ou mais morfemas distintos, realizados por uma única unidade, denominada na literatura **morfe portemanteau** (fr. 'cabide') ou **cumulativo**, como no caso da forma *du* do francês, que realiza dois morfemas distintos, a saber *de+le*.

²² A análise de *amas* não é a mesma. Não há aí, segundo Hockett (1947: 238), qualquer morfe vazio: nesse caso, /a/ indica IND. PRES., uma vez que é a única diferença em relação a *ames*, por exemplo (cf. seção 3.2.2.).

Em virtude dos problemas suscitados pela noção de morfema segundo o modelo IA, começariam a surgir propostas alternativas para lidar com a morfologia. A análise de Matthews (1972) sobre o verbo em latim²³ advogaria a volta do modelo Palavra e Paradigma no estudo das chamadas línguas flexivas. Era o começo de propostas voltadas para a palavra, como veremos a seguir.

²³ Na lingüística atual, talvez a primeira voz em favor da importância da *palavra* na análise gramatical seja Robins (1959).

4

Preparando o retorno da palavra

4.1. Introdução

No presente capítulo vamos focalizar alternativas à noção clássica de morfema nascidas nas últimas três décadas. Surgiram tanto no âmbito da morfologia derivacional como no da flexional, em conseqüência de problemas como os que acabamos de ver. Das muitas propostas, quase uma por autor, este capítulo tem como pano de fundo quatro, em virtude da repercussão desses trabalhos internacionalmente e, em especial, no Brasil: Matthews (1972; 1974), Aronoff (1976; 1994), Basílio (1980), Anderson (1985b; 1992). Para uma visão ampla das propostas surgidas nas últimas décadas, sugerimos Spencer (1991) e Carstairs-McCarthy (1992).

4.2. O morfema na derivação

Na morfologia por Item e Arranjo, os morfemas são as unidades que constroem a análise em CIs. Esses 'pedaços de material', em geral (ou pelo menos idealmente) em seqüência, juntam-se para formar sucessivamente palavras, sintagmas, orações, frases. Esse modo de focalizar o morfema está captado na metáfora 'o morfema é uma coisa' – *i.e.*, uma forma –, e a morfologia, por conseguinte, é afixação por excelência.

A definição do morfema como *forma mínima com significado* levou a dificuldades que não deixaram de ser detectadas pelos estudiosos da época, a começar pela questão do significado a atribuir a um morfema. Em seu clássico *Morphology*, Nida (1949: 162) afirmava que "os significados das formas presas são particularmente difíceis de descrever". Em especial as dificuldades em estabelecer com precisão o significado de um morfema se fizeram sentir no estudo da derivação (Basílio, 1974a: 85).

Vamos ilustrar tais dificuldades com dois conjuntos de exemplos recorrentes na literatura. Primeiramente, que significado atribuir a formas mínimas que reconhecemos como recorrentes, mas cujo significado nos escapa? Que fazer com exemplos como *re-ceb-er*, *con-ceb-er*, *de-duz-ir*, cujas raízes (e também os prefixos) parecem não ter qualquer significado?

Por outro lado, que fazer com os *hapax legomena* (gr. 'que foram ditos apenas uma vez'), ou seja, com formas para as quais se encontra somente um exemplo na língua? Em inglês, *cran-* aparece apenas na palavra inglesa *cranberry* 'uva-do-monte', e é isolável em resultado da comparação com formas terminadas em *-berry*, denominativas de pequenos frutos silvestres, como *strawberry* 'morango', *blueberry* 'fruto do vacínio', *blackberry* 'amora preta' e *gooseberry* 'groselha'. No entanto qual seria o significado de *cran-*? É bom lembrar que o primeiro dos seis passos para determinar o significado dos morfemas (Nida, 1949: 162 – ênfase no original) dizia respeito justamente ao número de dados:

1. Faça a coleta de muitas ocorrências de um morfema. É quase impossível chegar a uma definição de um morfema na base de duas ou três ocorrências. É claro, podem-se tentar hipóteses acerca do significado, mas no mínimo oito ou dez ocorrências deveriam ser checadas.

Decorre dessa estratégia o fato de que "o analista se vê diante de dois caminhos: a) não considerar tais formas [como *-ceb-*, *-duz-*, *-fer-*, *MCR*] como morfemas isolados, apesar da recorrência; b) não considerar a atribuição de significado como elemento decisivo para considerar ou não uma seqüência fonêmica como morfema" (Basílio, 1974a: 85). Qualquer dessas soluções é indesejável.

No estudo da formação de palavras, os problemas com os significados das formas não levou ao puro e simples abandono da noção de morfema, mas a um redimensionamento desse conceito. Retomemos parte dos exemplos acima, acrescentando-lhes alguns mais:

(4.1) referir	reduzir	receber
deferir	deduzir	
conferir	conduzir	conceber
inferir	induzir	
preferir		
transferir		

Podemos constatar a recorrência das raízes *-fer-*, *-duz-* e *-ceb-* nesses dados, mas conseguimos atribuir significado não a elas, mas tão-somente às palavras como um todo. No entanto, como nota Aronoff (1976: 12-13) a respeito de exemplos seme-

¹No mesmo caso de *cran-* estão *boysen-* em *boysenberry* 'tipo de amora' e *huckle-* em *huckleberry* 'mirtilo'. Formas desse tipo são referidas na literatura como *morfes cranberry*.

lhantes no inglês, nos nomes derivados *-cep-* toma sistematicamente o lugar de *-ceb-* (*conceber/concepção*, *receber/recepção*), num processo comum de alomorfa, não fora a questão do significado. Não se trata de atribuir *status* a qualquer seqüência fônica que se repete, mas de atribuir esse *status* a seqüências que se relacionam a uma entidade lingüística fora dela (Aronoff, 1976: 15). Para Aronoff (1976: 10ss) tais formas são morfemas, embora sem significado.

Mas não é só isso. Uma forma como *-vel*, por exemplo, presente em palavras como *dobrável*, *reversível*, não existe independentemente de uma **Regra de Formação de Palavras**² que forma adjetivos a partir de verbos em português com o significado 'que pode ser X-do' (onde X representa o verbo derivante), ou de uma regra que analisa³ adjetivos existentes língua.

Uma palavra pode mudar de significado com o tempo, e seu significado já não ser previsível a partir dos elementos que a constituem: *amável* 'gentil', *considerável* 'muito/grande' são alguns exemplos. Novas palavras em *-vel* apresentam o significado esperado, o que leva a considerar que foi a palavra como um todo, e não o sufixo, que sofreu mudança semântica.

O redimensionamento do conceito de morfema significou retirar dele o papel central na análise morfológica. Na busca de explicação para a **competência lexical** dos falantes, o que passa a ser necessário não é o estabelecimento de listas de elementos mínimos, mas a resposta a questões acerca de que palavras os falantes podem formar, que tipos de palavras, novas ou antigas na língua, são capazes de analisar, que relações estabelecem no âmbito do vocabulário. É esta, *grosso modo*, a visão de Aronoff (1976), de Aronoff & Anshen (1998) e de Basílio (1980). Os processos produtivos de formação de palavras (a questão central em Aronoff, 1976) atuam sobre palavras existentes na língua. Essa hipótese é conhecida como **morfologia baseada em palavras**. Para não ser facilmente falsificada, a noção de *palavra* teve de ser "ajustada", como veremos no próximo capítulo.

Preocupada tanto com as formações produtivas como com o reconhecimento da estrutura de formas existentes na língua que não resultam de processos produtivos, Basílio (1980) tem este como um dos pontos discordantes em relação a Aronoff (1976).

4.3. A flexão: o abandono do morfema

No que toca à flexão, os ataques ao morfema foram mais ferozes e levaram boa parte dos lingüistas que se dedicam à morfologia flexional, em especial no trabalho com línguas flexivas, a abandonar a noção de morfema. Isto não significa dizer que não há interesse em identificar as relações entre partes da forma de uma palavra e

²Denominação em Aronoff (1976).

³Para Aronoff (1976), a análise da estrutura é a contraparte da **Regra de Formação de Palavras**. Para Basílio (1980), isto é papel de outra regra, a **Regra de Análise de Estrutura**.

partes de seu significado. Essa é uma tarefa que continua sendo relevante; contudo, como nota Anderson (1985b: 160), “é necessário ir além da simples noção de que há uma correspondência de um-para-um – como implicado na noção de que o morfema é a unidade mínima de análise – que estabelece uma associação direta entre forma e significado”.

Um exemplo simples: a forma verbal *amo* pode ser segmentada em *am-o*, e poderíamos dizer (como é usual) que *-o* é o morfema de Primeira Pessoa/Singular. Sabemos, no entanto, que essa marca é “condicionada” pelo IND./PRES., isto é, que não temos outra 1SG em *-o* no verbo AMAR. Esse *-o*, portanto, marca Pessoa/Número, mas também Modo/Tempo. Propor um morfema modo-temporal zero para o IND./PRES. nada mais é que um artifício descritivo que procura manter o padrão canônico da estrutura do verbo em português. De qualquer modo, a presença desse *-o* nos diz que essa é a *primeira pessoa do singular do presente do indicativo* do verbo AMAR. O que queremos dizer com isso é que a relação entre *-o* e 1SG, e entre IND./PRES e \emptyset é mais complexa do que possa parecer nessa análise em morfemas. Por quê?

4.3.1. Um-para-um versus um-para-muitos, muitos-para-um

O morfema clássico captava nosso conhecimento de que determinadas seqüências fônicas expressam determinados significados. O grande problema foi focalizar de modo relativamente simples essa relação, transformando-a numa relação direta entre uma seqüência sonora e um significado e vice-versa, ou, como geralmente referida, como uma **relação um-para-um**. No trabalho com diferentes línguas do mundo multiplicaram-se os dados para os quais tal visão era insatisfatória.

Carstairs-McCarthy (1992: 194-195) aponta quatro tipos de afastamento do padrão *um-para-um*. Em primeiro lugar, uma propriedade flexional pode ter várias realizações numa única forma de palavra. É o que Carstairs-McCarthy classifica como **desvio I: um-para-muitos sintagmático**, que corresponde ao que Matthews (1972; 1974) denomina *exponência expandida*. A distinção entre o Presente e o Perfeito nas formas do grego *le-lu-k-a* ‘eu perdi’ e *lu-o* ‘eu perco’ está marcada: (a) na reduplicação ou *redobro (le)*; (b) no sufixo *-k-*; e (c) na forma que indica Primeira Pessoa/ Singular.

Uma propriedade pode realizar-se de várias maneiras, mas em palavras diferentes. Segundo Carstairs-McCarthy (1992), é o **desvio II: um-para-muitos paradigmático**. Em português, por exemplo, a Segunda Pessoa/ Singular realiza-se como *-ste* no Perfeito do Indicativo, como \emptyset no Imperativo, como *-s* nas demais formas.

Por outro lado, diferentes propriedades morfossintáticas podem realizar-se num único elemento da forma de palavra. Em *insul-īs*, *-īs* indica que a forma latina é, ao mesmo tempo, Ablativo e Plural. Para Carstairs-McCarthy (1992), é o **desvio III: muitos-para-um sintagmático**. Na classificação de Matthews (1972; 1974), é a *exponência cumulativa*.

Por fim, as relações de *muitos-para-um* podem ser um caso de homonímia flexional: *insul-īs*, por exemplo, podem representar o Ablativo Plural ou o Dativo Plural. Constituem-se no **desvio IV: muitos-para-um paradigmático**.

4.3.2. Morfemas versus formativos, expoentes

Se não queremos considerar que todas as relações sejam obrigatoriamente de um-para-um, que fazer? A solução alternativa foi a de distinguir dois níveis: (a) o nível da estrutura semântica, que envolve as *raízes* e as *categorias gramaticais* (vide cap.8); e (b) o nível da expressão fonológica, em que aparecem os processos gramaticais, tais como alternância nas raízes – seja vocálica, consonantal, de tom ou de acento –, a afixação, a reduplicação. Neste segundo nível temos os **formativos** ou, especificamente fazendo referência à expressão de uma categoria, os **expoentes** (Matthews, 1972; 1974). A diferença entre morfemas de um lado e formativos ou expoentes de outro pode parecer apenas de nomenclatura, mas não o é. Ela representa uma ruptura com a noção de que o morfema é um signo: o significado e sua expressão devem receber tratamento em separado.

4.4. Morfologia baseada em morfemas e morfologia baseada em palavras

Ao propormos (tal como Matthews, Anderson, Aronoff) que a morfologia apresenta dois níveis, um que lida com os *formativos* e outro cujo enfoque está no material sintático-semântico, estamos assumindo que o morfema não é a unidade básica de análise, mas que esse papel cabe à palavra. Segue-se daí que a morfologia e a sintaxe não são basicamente a mesma coisa, como a análise em constituintes imediatos levava a crer. Teorias em que a morfologia tem a palavra (no sentido de *lexema*) como sua unidade básica contrapõem-se, assim, às teorias que tomaram o morfema como a unidade básica. Essa distinção está marcada na nomenclatura: **morfologia baseada em palavras** ou **morfologia baseada em lexemas**, de um lado e, de outro, **morfologia baseada em morfemas**. Passemos agora a precisar o termo *palavra*.

5

Quantas palavras temos num enunciado?

5.1. Introdução

Até aqui usamos o termo *palavra* num sentido pré-teórico, isto é, sem defini-lo, enfim, sem enquadrá-lo no âmbito de uma teoria. Como *palavra* é um termo que existe no uso cotidiano da língua, pode parecer, a princípio, uma tarefa muito simples defini-lo. Não é bem assim.

De acordo com o uso comum do termo, que tem por base nosso conhecimento da escrita, parece trivial definir o que seja uma palavra. Qualquer criança no ensino fundamental já sabe responder quantas e quais são as palavras em *Penélope ama Odisseu*. Talvez já fique em dúvida perante *ama* e *amava*, não sabendo se as contará como uma ou como duas palavras, quando, então, a pergunta parece levar em conta algo que ultrapassa uma determinada seqüência de letras ou símbolos. Ou ainda se contará também como palavras elementos como *que*, *se*, *de*, quando entra em jogo o tipo de significado das formas.

Deixada de lado a escrita, porém, tem a palavra, afinal de contas, algum tipo de relevância teórica que a torne interessante para a pesquisa gramatical, ou, ao contrário, poderia ser considerada apenas um epifenômeno, isto é, um fenômeno secundário?

A resposta parece ser sim para a primeira parte da pergunta. A despeito do contínuo sonoro existente na oralidade, os falantes conseguem abstrair parte das características físicas de um enunciado e desenvolver, a partir da mais tenra infância, algum tipo de estratégia que lhes permite segmentar esse contínuo em unidades menores, com base no ritmo do *input* nativo, seja ele acentual, silábico ou moraico.

Pesquisas com línguas diferentes têm demonstrado que há estratégias que não são universais, mas dependentes das características rítmicas específicas da primeira língua, ou da língua dominante no caso de bilíngües, para a apreensão dessa unidade a que estamos denominando intuitivamente *palavra* (Cutler, 1994). Ao ouvirmos o contínuo sonoro, lançamos mão dessa estratégia, e ela é surpreendentemente eficaz mesmo quando os sinais sonoros da fala vêm misturados com outros sons do ambiente, ou pronunciados com sotaque estrangeiro, ou a uma boa distância do ouvinte.

Falantes do inglês como primeira língua, por exemplo, segmentam o enunciado com base no ritmo acentual entre sílabas fortes e fracas (Cutler, *id. et ibid.*). Uma vez que a maior parte das palavras em inglês tem sílaba tônica inicial, os erros de percep-

ção são mais comuns se produzem a inserção de uma fronteira de palavra diante de uma sílaba forte (como em *achieve* ouvido como *a cheap*) ou o apagamento de uma fronteira de palavra diante de uma sílaba átona (como no caso de *bird in* ouvido como *burgling*). São menos usuais quando produzem a inserção de uma fronteira diante de uma sílaba átona (como em *effective* ouvido como *effect of*) ou o apagamento de uma fronteira diante de uma sílaba tônica (como *were waiting* sendo ouvido como *awaiting*). Por seu turno, falantes de francês usam como estratégia de segmentação o reconhecimento *sílaba a sílaba*¹, ao passo que falantes de japonês usam como estratégia a *mora* (Cutler, 1994: 92-93)².

Para a lingüística, o grande problema em definir *palavra* é ser esse termo passível de receber diferentes caracterizações nas diferentes dimensões do estudo da linguagem, nem sempre resultantes na mesma unidade. Afora o uso na escrita, podemos entender *palavra*: (a) como uma unidade fonológica; (b) como o elemento mínimo da estrutura sintática; (c) como um elemento do vocabulário da língua. Nas secções que se seguem vamos focalizar essas diferentes interpretações. A começar pela mais usual. Vamo-nos deter um pouco nas justificativas para não levar em conta, aqui, a *palavra gráfica*.

5.2. A delimitação da palavra

5.2.1. A palavra gráfica

Utilizamos o termo *palavra*, cotidianamente, como uma noção oriunda da escrita, que se aplica também à fala. Por que motivo essa caracterização não será aqui privilegiada? Simples: porque consideramos a escrita como parte de um outro estudo, o dos recursos expressivos particulares que a representação gráfica torna disponíveis (Nunberg, 1990: 7). A escrita tem sua própria "gramática", isto é, seu conjunto particular de regras.

Aprender a escrever é, em parte, aprender a identificar e a assinalar os limites da **palavra gráfica**, unidade delimitada por **separadores**, *i.e.*, por espaços em branco ou quebras de linha, mas também por sinais de pontuação ou ainda por letras de traçado diferenciado, consoante ocupem ou não a posição final na palavra escrita³.

Embora nas escritas modernas ocidentais que empregam o alfabeto latino a no-

¹ O ritmo do português do Brasil estaria em mudança: de um padrão silábico como o francês ou o espanhol, para o acentual, como o inglês (Major, 1981).

² Que pode ser uma parte da sílaba que não o *onset*.

³ Vimos exemplo deste último tipo nas primeiras aulas de Grego, quando aprendemos a distinguir o sigma final (ς) do não-final (σ).

ção de *palavra gráfica* seja, no mais das vezes, equivalente a uma forma livre, ela é tratada por regras da escrita. Talvez onde melhor possamos percebê-lo, no caso do português moderno, seja na representação de compostos, de preposições e de clíticos. Com a ortografia atual para o português do Brasil por base, por exemplo, temos de considerar fatores como os hifens ao definir uma palavra composta: *pé de anjo* ('aquele que tem pé grande'), por exemplo, é uma seqüência composta por três palavras gráficas, mas *pé-de-cana* ('aquele que bebe muito') é apenas uma palavra, pelo menos até o aparecimento de alguma reforma ortográfica que venha a banir os hifens, ou que, ao contrário, inflacione seu uso, quando, então, o quadro passaria a ser outro. Nosso estudo de morfologia teria de incluir um capítulo sobre o uso de hifens? Seria estranho, uma vez que o uso de hifens é regulamentado, em última análise, por decreto do presidente da República...

Não podemos esquecer que nossa escrita resulta de cerca de quinhentos anos de padronização, que começou a ser efetivamente imposta a partir da invenção, no Ocidente, da imprensa de tipos móveis. A palavra gráfica, como a conhecemos, é um artifício relativamente recente, e escritas mais antigas nem sequer segmentaram o espaço do suporte⁴. A chamada *escrita bustrofedônica* dos gregos (século VI a.C.), por exemplo, fazia-se em linhas horizontais, alternadamente, da direita para a esquerda e vice-versa, algumas vezes de baixo para cima na superfície escrita (Diringer, s.d.: 148), sem qualquer separação entre palavras.

Num tipo de escrita diferente da nossa, a hieroglífica do antigo Egito (ca. 5000 a.C. a 100 d.C.), a fronteira gráfica que delimitava o final de uma palavra era marcada por um símbolo especial, o *determinativo*, que podia combinar-se (ou não) com até dois outros. Os 180 determinativos tinham por função classificar as palavras: o rolo de papiro selado () , por exemplo, indicava que aquela palavra pertencia à classe das idéias abstratas (Jacq, 1994: 27). Seria uma nova palavra o que viesse antes do determinativo, ou depois dele, dependendo da direção para a qual as figuras humanas ou de animais se voltavam no suporte. Para os nomes de reis a visualização era mais fácil: o *cartucho*, uma linha ovalada, envolvia o conjunto de símbolos que compunham a palavra, assim, , e sinalizava qual o conjunto a levar em conta em meio a uma seqüência sem solução de continuidade.

Para as escritas alfabéticas, talvez possamos considerar os primeiros passos na direção da *palavra gráfica* o desenvolvimento de formatos diferenciados para alguns grafemas, o que viria a ajudar o processamento da leitura de um conjunto compacto de linhas e, nas inscrições monumentais romanas (séculos I e II), o uso de um ponto entre palavras.

Mesmo depois de os separadores começarem a ser empregados, seu uso não

⁴ Dá-se o nome de **suporte** ao material em que um texto está inscrito. Esse material tem variado ao longo dos séculos: pedra, papiro, pergaminho, papel, vinil, disco magnético, película fotográfica.

coincidiu de imediato com o das escritas modernas. Na Europa medieval, preposições e palavras curtas foram, em geral, unidas à palavra seguinte (Bischoff, 1986: 173). Mas, pelo menos até o início do século XVI, obras manuscritas e impressas estiveram repletas de formatos diferenciados para os grafemas, consoante sua posição na palavra, e de *conglomerados gráficos* inesperados pela óptica de um leitor moderno, os quais, por vezes, induziram até mesmo renomados filólogos a erros de leitura.



Um exemplo famoso de erro de leitura, causado por separadores utilizados de forma diferente daquela a que nos habituamos, originou-se de um verso da égloga *Crisfal* (Cristóvão Falcão, 15-): *cantou canto de ledino*. Teófilo Braga (1875), no *Manual da história da literatura portuguesa*, interpretou esses desconhecidos *cantos de ledino* como '*cantos de romaria*', explicando que eram cantos alegres porque, nas romarias, as moças tinham ocasião de encontrar-se com seus amados. O texto de Braga influenciaria outros filólogos, como E. Monaci, que lhe dedicaria toda uma obra (*Cantos de ledino tratti dal grande Canzoniere portoghese della Biblioteca Vaticana*, 1875).

Carolina Michaëlis de Vasconcelos viria mais tarde a demonstrar que a leitura correta do verso seria: *cantou canto d'ele dino*. Dito de outra forma: *cantou um canto digno dele*. (Para mais detalhes, Roncaglia, 1974-5: 88-89).

Os separadores em textos antigos podiam, mesmo, refletir um estágio da língua que já então era passado: formas como *tal vez*, *por tanto* e advérbios em *-mente*, por exemplo, continuavam a ser grafadas como duas palavras gráficas tempos depois de já se terem combinado numa só. A título de ilustração, veja-se o excerto abaixo, extraído da *Estoria de muy noble Vespesiano emperador de roma* (Anônimo, 1496. fol. c3^{r-v}), um dos romances do ciclo do Graal:

E pilatus & el rey archileus com dez caualleyros. se sobirõ no muro sem armas. & vestidos de senhos briaes vermelhos [...] O nobre meu padre te encomẽdou esta çidade por que aguardarres & arregarres por elle [...] E de pois de sua morte enuiar te me o trebuto [...] E [...] derprezar teo muyto mal.

Nesse pequeno exemplo, os pronomes átonos ora estão, ora não, formando uma unidade gráfica com o verbo; o advérbio *depois* (< lat. *de post*) aparece como duas palavras gráficas, revelando a consciência do étimo latino. E que fazer quando nos deparamos com termos desconhecidos, como o distributivo⁵ *de senhos*?⁶. Observe-se ainda o traçado diferente para <s>, em conformidade com sua posição na palavra gráfica.

Em suma: a função básica da (orto)grafia é o registro da informação de modo

⁵ Dá-se o nome **distributivo** a um item que refere individualmente cada elemento de um conjunto, como *cada, todo, nenhum*.

⁶ *De senhos* (< lat. *singuli*) significa 'cada um'.

permanente (*verba volant, scripta manent*, já dizia o antigo provérbio) e não a transcrição sistemática e inequívoca de dados para o trabalho de um linguista⁷. Nem se imagine identificar os separadores com as pausas na fala. Quando falamos, não fazemos pausas a cada forma livre ou dependente⁸, mas entre seqüências com extensão média de cinco a seis palavras (Chafe, 1992: 25). Deixemos então de lado a escrita e passemos à fala.

5.2.2. A palavra fonológica

O que é denominado **palavra fonológica** – unidade formada por fonemas, sílabas e traços supra-segmentais – pode ser menor do que aquilo que desejamos considerar uma palavra no estudo morfológico. É a um afixo, considerado ou não como uma palavra fonológica, que recorre a análise para distinguir, em italiano: (a) prefixos produtivos que, quando se agregam a uma palavra começada por /s/, não configuram ambiente para a sonorização da fricativa, como em (5.1)⁹; de (b) formações improdutivas que permitem a sonorização de /s/, como em (5.2). No primeiro caso temos duas palavras fonológicas¹⁰, uma que corresponde ao prefixo, outra, à base:

(5.1)	(ri) _p (suonare) _p	[risuonare]	'soar novamente'
	(a) _p (sociale) _p	[asoçiale]	'não social'
	(pre) _p (sentire) _p	[presentire]	'ouvir antecipadamente'

(5.2)	(presentire) _p	[prezentire]	'ter pressentimento'
-------	---------------------------	--------------	----------------------

De outro modo, em italiano, no interior de uma palavra fonológica, o ambiente intervocálico impede o aparecimento de /s/ não geminado:

(5.3)	(rosa) _p	[roza]
	(cosa) _p	[coza]

Em português também há exemplos em que podemos considerar a palavra fonológica menor que a forma livre. É o caso de derivados em *-mente*, *-íssimo* e *-inho*, que podemos considerar como constituídos de duas palavras fonológicas. A postulação de uma fronteira entre palavras fonológicas explicaria o não fechamento da vogal

⁷ Aliás, foi essa mesma conclusão que levou ao estabelecimento de alfabetos fonéticos.

⁸ Quando muito, os espaços em branco da escrita poderiam corresponder a *pausas potenciais*. Se enunciada bem lentamente, no entanto, uma frase poderia apresentar pausas também entre suas sílabas: *a/ bo/me/cal/ del/ pa/no es/tá/ su/ja*.

⁹ Exemplos extraídos de Spencer (1996: 178-179), cuja análise reporta-se a Nespor, M. & Vogel, I. 1986. *Prosodic Phonology*. Dordrecht: Foris.

¹⁰ Indica 'palavra fonológica'.

média pretônica no derivado, correspondente à tônica no termo derivante, uma exceção ao fato de, no dialeto carioca, as vogais médias abertas não ocorrerem em ambiente pretônico:

(5.4) *forma* ['foxma], mas *formoso* [foxmozu]

Do mesmo modo *clero*, com [ɛ] faz o derivado *clerical*, com [e]. A fronteira entre palavras fonológicas, no entanto, previne o fechamento da vogal média:

(5.5) (*seri*)_p (*íssimo*)_p (*nova*)_p (*mente*)_p (*pε*)_p (*zinho*)_p

Ao contrário desses exemplos do italiano e do português, há *palavras fonológicas* que não gostaríamos de considerar como uma única palavra num estudo morfológico. É o caso de formas fonologicamente dependentes que se agregam a outros elementos da sentença como se fossem sílabas iniciais ou finais. Uma seqüência como *disse-lhe*, foneticamente ['dʒisiɫi] na fala carioca, serve para ilustrar o problema. Formas átonas como o pronome *lhe* do exemplo agregam-se fonologicamente a outra palavra do enunciado a eles contígua e recebem a denominação geral de **clíticos** (do gr. *klitikós*, 'que se inclina ou apóia' pelo lat. *cliticus*).

Um *clítico* é uma *forma dependente*. A classificação *forma dependente* foi criada por Mattoso Camara Jr. como complemento à distinção bloomfieldiana entre *formas livres* e *formas presas*. O problema classificatório não era novo. Também Sapir constatara, ao tratar do paiute¹¹, que havia formas que nem eram verdadeira afixação, nem se constituíam na justaposição de elementos independentes¹².



O próprio Bloomfield (1926: 26) reconhecera a existência de formas de caráter misto, e classificou-as como *formativos de sintagma*:

Pressuposto S1. Um sintagma pode conter uma forma presa que não é parte de uma palavra. Por exemplo, o possessivo [z] em *the man I saw yesterday's daughter*: ['a filha do homem que eu vi ontem' – MCR]

Def. Essa forma presa é um *formativo de sintagma*.¹³

¹¹ Língua uto-asteca do N de Nevada, Oregon, Califórnia e Idaho (EUA).

¹² "Enclisis is neither true suffixation nor juxtaposition of independent elements. It has the external characteristics of the former (including strict adherence to certain principles of order), the inner feeling of the latter." (Edward Sapir, 1930. *Southern Paiute, a Shoshonean language*. Apud Klavans, 1982: 1).

E completava: "Esse pressuposto perturba a definição de *sintagma* [...]"¹³ (Bloomfield, *id. et ibid.*).

A questão reapareceria em Nida (1949: 104), já então com a preocupação em justificar por que incluir o genitivo 's do inglês entre os afixos e não entre os clíticos¹⁴. Segundo Nida, os clíticos têm muito maior liberdade que afixos, e *s* ocorre apenas com nomes, pronomes e sintagmas nominais.

Os *formativos de sintagma* ou *afixos de sintagmas* formam, porém, um conjunto mais restrito que as *formas dependentes* de Camara Jr., porque estas incluem artigos e preposições, na medida em que, na fala comum, pertencem ao mesmo grupo acentual do núcleo do sintagma.

Os *clíticos* prendem-se fonologicamente a outra palavra no enunciado, que é o seu **hospedeiro** (tradução do inglês *host*)¹⁵. No português europeu, por exemplo, os clíticos de acusativo "são sempre fonologicamente enclíticos, a despeito de qual seja a palavra precedente", o que os impede de estarem no início absoluto da sentença (Nunes, 1992: 5, citando Carvalho, 1989¹⁶). O sinal = nos exemplos abaixo indica a cliticização:

- (5.6) a. *Quem=me vê?*
 b. *Já=te digo*
 c. *Não=te vi*
 d. *João tinha=me dado um livro*
 e. **Me=diga uma coisa*

Afora os clíticos, é interessante observar ainda que a fala espontânea apresenta muitas vezes o "apagamento" de fronteiras entre palavras. Uma seqüência como as

¹³ Bloomfield (1926: 27): "Uma forma livre não-mínima é um sintagma. E.g., *the book*. or *The man beat the dog*; mas não, e.g., *book on* [...], porque é sem significado e, logo, não é uma forma; tampouco *blackbird*, que é uma forma livre mínima."

¹⁴ Nida (1949: 104): "The genitive -s may occur with single morphological units, as in John's, his, and everybody's, or it may occur with phrases, e.g. the old man's (aches and pains) and the king of England's (hat). This morpheme is considered a suffix rather than a clitic because its distribution is limited to occurrence with nouns and pronouns and because the resultant construction belongs to one of two external distribution classes: (1) attributive to nouns, e.g. the king of England's (hat) or (2) a substitute for nouns, e.g. the king of England's was there. [...] The distributional behavior of the genitive -s is quite different from that of the usual clitic, which may be combined with any number of classes of forms and with numerous resultant external distribution class memberships. Ainda no âmbito do estruturalismo norte-americano, a mesma questão seria outras vezes retomada, como, e.g., em Wells (1947: 196ss).

¹⁵ Na dependência da posição que ocupam em relação ao hospedeiro podem ser *proclíticos*, *mesoclíticos* ou *enclíticos*.

¹⁶ Carvalho, J., 1989. Phonological conditions on Portuguese clitic placement: on syntactic evidence for stress and rhythmical patterns. *Linguistics*. 29, 405-436.

amigas [ðza'migðš] é um contínuo sonoro, em que a sílaba travada¹⁷ do artigo desaparece em razão de uma vogal iniciar a palavra seguinte, transformando uma seqüência VC#V em V-CV¹⁸. Se a *juntura*, por um lado, indica que não houve solução de continuidade na emissão da voz, por outro é ela um fenômeno típico de fronteira entre palavras em português, e que ajuda, por conseguinte, no reconhecimento dessa unidade. Os acentos (representados pelo sinal ^) indicam a presença de um nome, verbo, adjetivo, advérbio, que constituem, em geral, unidades também no nível fonológico. Em algumas línguas, mais precisamente aquelas que apresentam o *acento culminativo*¹⁹, como o latim, o checo ou o polonês, a localização do acento é fixa, servindo de auxiliar na identificação de palavras²⁰.

Poremos de lado a *palavra fonológica* neste trabalho. A existência de clíticos, de *juntura*, de regras como a do italiano a que aludimos anteriormente faz com que essa unidade não coincida, necessariamente, com o que gostaríamos de considerar palavra num estudo morfológico. Em princípio queremos que as formas livres e as formas dependentes façam parte do que entendemos como palavra. Assim, seqüências como *os meninos* ou *disse-lhe* serão aqui consideradas como compostas de duas palavras; o italiano *asociale*, por sua vez, como uma única.

Aliás, em estudo recente, Hanlon & Edmondson (1996) sugerem, com base em estudo sobre a afasia de jargão fonêmica²¹, que palavras fonológicas podem manter-se intactas apesar de o paciente não apresentar qualquer outra evidência de que essas unidades tenham significado ou classe gramatical, por exemplo²².

5.2.3. A palavra como unidade sintática mínima

Considerar a palavra como uma unidade mínima para a sintaxe significa dizer que a sintaxe não forma palavras. É o que a literatura gerativa denomina **hipótese lexicalista**. A sintaxe forma constituintes, frases, mas não palavras. Dito de outro modo: a estrutura interna da palavra não é da alçada da sintaxe, mas da morfologia.

¹⁷ Aquela que termina em consoante, representada como (C)(C)VC.

¹⁸ Em que C representa 'consoante' e V, 'vogal'; -, 'fronteira silábica', e #, 'fronteira de palavra'.

¹⁹ Opõe-se ao *acento livre* ou *distintivo* de línguas como o português, por exemplo.

²⁰ Em latim clássico o acento é determinado pela quantidade: em dissílabos acentua-se a penúltima sílaba. Em palavras de três ou mais sílabas, acentua-se a penúltima se esta for longa (ou por apresentar uma vogal longa, ou um ditongo, ou uma vogal breve seguida de duas ou mais consoantes); se for breve, acentua-se a antepenúltima.

²¹ Afasia em que o paciente produz seqüências fluentes de fonemas, mas destituídas de qualquer significado. O termo *afasia* foi utilizado primeiramente por Platão, para designar o silêncio daquele que se calava diante de um argumento definitivo, e seria retomado por Armand Trousseau em 1865 com o sentido de déficits lingüísticos (Guindaste, 1996).

²² Um pequeno exemplo: como resposta à pergunta "Diga-me como dormiu a noite passada, OK?", a paciente responde: 'hæ?e 'd,elɔn 't'ræ?n 't'ræk 'nɔzɔ.

Não é importante para a sintaxe que *reler* seja formado a partir de *ler* pela adjunção do prefixo *re-*²³. Daria no mesmo se estivéssemos diante da forma primitiva *ler*. Importa que *ler* (ou *reler*) seja verbo, uma vez que essa informação é relevante para fenômenos como concordância e regência. A hipótese lexicalista representa o reconhecimento de que as construções morfológicas são reguladas diferentemente da frase. Uma palavra e uma frase não são diferentes porque há mais unidades constituintes nesta que naquela, mas porque têm tipos de coesão interna diferentes.

Numa palavra os constituintes internos não têm mobilidade. A posição de um sufixo não pode ser preenchida por uma raiz; tampouco por um prefixo. Em línguas como o suaíli, que refletem pela concordância o Sujeito e o Objeto na estrutura do verbo, essa ordem interna não tem a mobilidade existente para S(ujeito), V(erbo) e O(bjeto) sintáticos: os formativos têm de apresentar-se na ordem [Sujeito – Tempo – Objeto -raiz]_v. Assim, a forma verbal que poderíamos traduzir como 'eu lhe darei' é aquela em (5.7):

(5.7) [ni - ta - m - pa]_v
1SG/S - FUT- 3SG/O - dar (raiz)

As unidades da sintaxe têm mobilidade incomparavelmente maior. Numa das chamadas *línguas de ordem livre*, como o latim, por exemplo, o número de posições possíveis para os elementos sintáticos é grande. Uma frase como 'Pedro vê Paulo' pode ser traduzida como (a), ou, por questões de ênfase, como (b), (c) ou (d) abaixo (baseado em Moreland & Fleischer, 1977: 5):

- (5.8) a. Petrus Paulum videt (ordem neutra).
b. Petrus videt Paulum (ênfase no sujeito e no objeto).
c. Paulum Petrus videt ('É Paulo que Pedro vê').
d. Videt Paulum Petrus ('Pedro realmente vê Paulo').

Uma palavra tem expansão limitada e regulada a partir de certo ponto, ao contrário da possibilidade de expansão infinita dos grupos de palavras. Por exemplo, as regras do português que formam derivados em *-al*, *-iz(a(r))*, *-ção* podem aplicar-se em seqüência. Assim, podemos derivar de *instituír*, *instituíção*, e, sucessivamente, em *camadas*, *institucional*, *institucionalizar*, *institucionalização*, ??? *institucionalizacional*. A expansão é possível, mas, a partir de determinado ponto, as

²³ Os hifens que representam a posição de uma forma presa em relação aos outros constituintes da palavra. Por exemplo: *des-*, *-ção*, *-duz-* indicam, respectivamente: (a) que a forma *des-* precede outra forma; (b) que *-ção* se segue a algo; (c) que *-duz-* (como em *deduzir*) necessita de um elemento anteposto a ele e de outro, que o siga.

derivações param. A expansão de uma unidade sintática, por sua vez, é, teoricamente, ilimitada: *o meu grande amigo, o meu grande e querido amigo, o meu grande, querido... e inesquecível amigo.*

① A hipótese lexicalista veio substituir a chamada **hipótese transformacionalista** da gramática gerativa, exemplarmente ilustrada em Lees (*The grammar of English nominalization*, 1960). A hipótese transformacionalista tratava processos gerais de formação de palavras, como a formação de nomes a partir de verbos como um processo sintático, as *nominalizações*: a mesma estrutura sintática geraria *João não quer que Pedro participe da reunião* e *João quer a participação de Pedro na reunião* (exemplos em Basílio, 1980: 26). A diferença entre as duas sentenças resultaria da aplicação de uma regra sintática (a *transformação*) de nominalização no caso da segunda.

Qual a vantagem que essa análise apresentava? A vantagem estava em captar generalizações, como a relação semântica existente entre um verbo e seu derivado nominal²⁴.

Se, no entanto, a estrutura interna da palavra é opaca para a sintaxe, que dizer dos expoentes para as *categorias morfossintáticas*? O Caso e o Número em latim, por exemplo, fazem parte da morfologia dos nomes e são relevantes para a sintaxe. Não há como formar uma oração correta em latim sem esse tipo de informação. O que nos leva a ver essas marcas de Caso e de Número como um tipo de morfologia diferente daquele que temos em formas como *-dade*. Podemos entender que a formação de palavras está em interação com o léxico; a flexão, por sua vez, representa a interação entre a morfologia e a sintaxe, ou morfossintaxe. Por conseguinte, nossa asserção acerca de a sintaxe ser cega para a estrutura da palavra, para não ser facilmente falsificada, necessita de um “ajuste” que distinga os processos flexionais daqueles que formam vocabulário na língua. Se tratamos a flexão e a derivação como tipos distintos de morfologia, podemos considerar seus elementos como entidades de caráter também diverso. Foi o que começamos a apresentar no capítulo 3, ao introduzirmos conceitos como *formativos* e *expoentes*. Os capítulos subseqüentes voltarão a essas questões.

5.2.4. A palavra como unidade da morfologia

Como já ficou evidente, o termo *palavra* é ambíguo. Mas se se ficarmos restritos apenas ao seu uso na morfologia? Essa ambigüidade persistiria? A resposta é ‘sim’. Vamos distinguir aqui diferentes significados que o termo *palavra* pode apresentar

²⁴ “Como princípios metodológicos” – explica Dillinger (1991: 58) – “preferem-se as generalizações e a eliminação de idiosincrasias. Conseqüentemente, uma das questões centrais da teoria lingüística é: até que ponto informações podem ser eliminadas do léxico sem sacrificar a adequação descritiva?”.

num estudo morfológico: como sinônimo de *lexema*; como sinônimo de *forma de palavra*; como sinônimo de *palavra morfossintática* ou *gramatical* (vide Matthews, 1972; 1974, 1991).

5.2.4.1. A forma de palavra

A **forma de palavra** é composta de uma seqüência sonora. A *forma de palavra* é o uso mais aproximado dos conceitos estruturalistas de *forma livre* (Bloomfield, 1926: 27) e de *forma dependente* (Camara Jr., 1970: 60).

Fazemos referência à forma de palavra quando dizemos estar diante de palavras homófonas: no português padrão do Brasil²⁵ *amamos* pode representar a *primeira pessoa do plural do presente do indicativo* de AMAR, mas também a *primeira pessoa do plural pretérito perfeito do indicativo* de AMAR.

5.2.4.2. O lexema

O **lexema** é uma palavra considerada como unidade abstrata. Tem significado lexical e pode apresentar variações, caso se inclua entre as *palavras variáveis*. O lexema pertence a uma das classes abertas da língua. Seguimos aqui a convenção de representá-lo com todas as letras maiúsculas.

Em geral, por motivos didáticos, traça-se um paralelo entre um **lexema** (parte do conhecimento de um indivíduo acerca de sua língua) e uma **forma de citação** (unidade de um dicionário em papel, ou na tela de um computador). Quando, na linguagem do dia-a-dia, referimo-nos a um verbo e não a uma de suas formas específicas, empregamos uma *forma de citação*: procuramos no dicionário o verbo *amar*, ou o verbo *vender*, ou *partir*. Se fosse um verbo latino, a consulta buscaria *amo*, *vendo*, *partior*, por exemplo. Com as formas de citação não fazemos referência somente ao infinitivo não flexionado português ou à primeira pessoa do presente do indicativo ativo latino, mas a todos os tempos, modos e pessoas desses verbos. A forma de citação *amar*, por exemplo, é virtualmente um conjunto de formas e nenhuma delas. O lexema é uma abstração nesse sentido: o lexema AMAR, ou SABER, por exemplo, representa a combinação virtual dos radicais que pode apresentar com todas as propriedades morfossintáticas com que se pode combinar.

5.2.4.3. A palavra morfossintática

E se quisermos falar não do verbo AMAR em geral, mas de um dos membros que formam seu paradigma? Por exemplo: se quisermos focalizar apenas a *primeira pessoa do singular do presente do indicativo*? Estaremos, então, com o mesmo AMAR,

²⁵ No português de Portugal, a abertura da vogal tônica distingue as duas formas. Já no português não-padrão do Brasil, o presente *amamos* distingue-se do pretérito *amamos*.

mas agora acompanhado de uma única entre as possibilidades de realização das categorias gramaticais ou morfossintáticas Tempo/Modo/Aspecto e Número/Pessoa. A **palavra gramatical** ou **morfossintática** é o lexema (AMAR, por exemplo) mais determinadas propriedades morfossintáticas, como IND/PRES/ISG.

Nos próximos capítulos nossa atenção estará voltada para o lexema e para a palavra morfossintática. E também para um conceito relevante para os estudos morfológicos: o *léxico*.

Súmula

Os problemas com a análise morfêmica clássica são variados e derivam, basicamente, do modelo de análise adotado, IA. Os problemas são interdependentes e alcançam até mesmo, como aponta Basílio (1974a: 82),

a não abrangência da metodologia, se considerarmos como morfes apenas os elementos aos quais podemos atribuir significado; ou se resumem na não adequação da metodologia às definições básicas de que ela partiu, se considerarmos também como morfes os elementos mínimos que restam após a segmentação em morfes.

Ao se desmembrar, no estudo da morfologia, significado e expressão, abria-se caminho para o retorno da palavra ao cenário, mas não mais como uma forma (*i.e.*, “um traço vocal recorrente que tem significado” – Bloomfield, 1926: 27) livre ou dependente. Por que não? Porque, ao tratar a palavra como uma unidade que se concretiza num enunciado, ela foi definida por um critério de pouco interesse para os estudos morfológicos: se pode ou não ser enunciada sozinha. Para o estudo morfológico, as características gramaticais são mais interessantes do que essa.

Em suma: a palavra é “por definição, uma abstração gramatical” e “qualquer critério fonológico deve permanecer logicamente secundário” (Robins, 1959: 120). O *lexema* e a *palavra morfossintática* são os elementos que nos interessam aqui. Com eles podemos dar conta das relações existentes no vocabulário da língua como também da informação gramatical que carregam.

III

REVISITANDO AS PARTES DO DISCURSO

PALAVRAS-CHAVES

morfologia e léxico;
flexão e derivação;
item lexical;
Regra de Formação de Palavras e produtividade;
bloqueio;
classe gramatical/parte do discurso;
significado lexical/significado gramatical;
classes abertas/classes fechadas;
categorias e propriedades.