

Coleção Espiral Volume 05 Série Lingüística

Luiz Carlos Cagliari

QUESTÕES DE
MORFOLOGIA E FONOLOGIA

Edição do Autor
Campinas, SP
2002

© - 2002 **Luiz Carlos Cagliari**

Todos os direitos autorais desta obra estão reservados ao autor, na forma da lei. Esta obra não poderá ser reproduzida em parte, nem no todo, por nenhum meio, sem a devida autorização do autor.

Esta obra contou com a colaboração do CNPq, em 1999 e 2000, através da Bolsa de Produtividade em Pesquisa (1-A), relativa ao processo N° 301450/78-1.

Capa de Daniel Raizer Cardinali Cagliari
Foto 2 – Tele/fax (0xx19) 3234-0150

C131q

Cagliari, Luiz Carlos, 1945
Questões de morfologia e fonologia / Luiz Carlos
Cagliari. - Campinas: L. C. Cagliari, 2002.
222 p.; 21 x 15 cm. – (Espiral, Lingüística ; v. 5)

ISBN: 85-900375-6-8

1. Morfologia. 2. Fonologia. I. Título. II. Série.

CDD-415

Para Gladis, Gianluca
e Daniel

Prefácio

Este é o quinto volume de uma série de livros que foram publicados pelo autor, como forma de difusão de apontamentos de aula e de trabalhos de pesquisa que, de outra forma, não estariam publicados.

O objetivo deste, como dos demais livros, é o de ajudar alunos de graduação e de pós-graduação a fazer leituras preliminares de introdução a questões de fonética, fonologia e morfologia. Neste quinto volume, é dada uma atenção especial a questões de morfologia relacionadas sobretudo com a fonologia.

Dado o caráter didático da obra, muitas informações gerais e introdutórias aos tópicos tratados são apresentadas, sendo, obviamente, dispensáveis aos leitores já familiarizados com elas. Para alguns, talvez, sirvam de revisão e de re-organização de noções básicas num quadro mais geral e abrangente.

O título original do trabalho era *A forma fonológica de base*, e tinha como objetivo a apresentação e a discussão de apenas esse tópico. Porém, dadas as dificuldades de apresentação de uma discussão direta desse assunto, sem remeter o leitor menos preparado com relação a outras informações, tidas como pressupostas, o livro acabou juntando os dois tipos de material, em grande parte mesclando-os.

A importância da Teoria da Otimalidade nos dias de hoje motivou a apresentação de uma visão geral introdutória dessa teoria e considerações a respeito da especificação da forma dos *inputs*.

Agradeço a colaboração recebida de colegas, sobretudo de minha mulher, Gladis Massini-Cagliari. Incoerências, inconsistências, falhas, lapsos e limitações, são apenas de minha responsabilidade. Mas, sem o incentivo dos amigos, esta obra não teria sido feita.

Campinas, 10 de Janeiro de 2002

Luiz Carlos Cagliari
E-mail: cagliari@lexxa.com.br

PRIMEIRA PARTE

Primeiras Idéias

1. Introdução

As primeiras manifestações de escrita de um texto, feitas pelos sumérios, por volta de 3300 a.C., representavam palavras inteiras (ideogramas, glifos). Tratava-se de uma segmentação fácil, de certo modo, uma vez que as palavras sumérias eram formadas tipicamente por monossílabos. Unindo monossílabos, podiam ter palavras compostas, como acontecia com os números. Neste caso, a representação das unidades aritméticas mostrava claramente que palavras compostas de outro tipo, por exemplos, formações novas por composição, empréstimos estrangeiros, nomes próprios, etc., também podiam ser analisadas em partes menores, embora o significado total dos compostos nem sempre representasse a soma dos significados das partes, como acontecia com a questão aritmética. Isso se tornava mais evidente, quando se tratava de nomes próprios ou de palavras estrangeiras, usados de forma adaptada à língua suméria.

O mesmo tipo de fenômeno é encontrado, ainda hoje, na língua chinesa. Uma palavra como Rio de Janeiro é pronunciada em chinês [lǐ-yě-ǎn-nèi-lú]. Ao interpretar os ideogramas correspondentes, nota-se que eles possuem significados próprios na língua chinesa, correspondentes a *cidade - concorda - quente - dentro - braseiro* - o que pode estar, de certo modo, associado às características da cidade do Rio de Janeiro, mas não é a simples tradução de *Rio de Janeiro*. A escrita foi feita com vistas à pronúncia de quem lê e não ao significado, como acontece, em geral, com a escrita ideográfica aplicada à maioria das palavras da língua. Apesar disso, a referência à idéia de “cidade” e não de “Rio”, mostra uma certa preocupação típica dos sistemas ideográficos.

A escrita obrigou os falantes a refletirem mais sobre a própria língua de uma maneira explícita e consciente, ao contrário da intuição que os falantes nativos têm ao aprender e ao usar a linguagem oral. Isso não significa que a linguagem escrita seja mais complexa, difícil ou mais perfeita do que a linguagem oral. Pelo contrário, tudo acontece primeiro na linguagem oral e, somente depois, por representação em outro meio de comunicação. Desta forma, a escrita

adquire algum sentido somente quando encontra por trás a expressão da linguagem oral.

Na sua essência, a linguagem é a soma de significados e significantes ou de idéias e sons, como bem foi definida desde os estruturalistas até o gerativistas. Os sumérios pegavam as palavras pelo significado e as representavam com pictogramas, no início da produção de sua escrita, os quais se transformaram em ideogramas cuneiformes, com o passar do tempo. Palavras ambíguas, como *ti* que significava por um lado *flecha* e, por outro, *vida*, não precisavam de dois ideogramas diferentes, porque o falante nativo sabia distinguir o uso de um do uso do outro, reconhecendo os diferentes significados pelos contextos em que ocorriam. Isso facilitava muito a escrita de palavras para as quais não se achava um pictograma de fácil representação gráfica. Palavras estrangeiras (por exemplo, nomes próprios) sofriam uma adaptação fonológica e morfológica, não semântica, que era natural para o usuário da linguagem, uma vez que essas palavras pertenciam a uma categoria semântica muito peculiar. Porém, para o usuário da escrita, a adaptação levava a considerações semânticas, derivadas dos ideogramas, que, em geral, eram bizarras, como no exemplo do nome da cidade do Rio de Janeiro, em chinês. Interpretações estranhas com segmentações impróprias, principalmente de nomes estrangeiros, costumam fazer parte de uma tradição jocosa, quase sempre preconceituosa, de uso das línguas. Nesse sentido, muitos nomes japoneses são usados por falantes de Português, para motivar piadas.

Deixando de lado outros aspectos da estrutura das línguas, nota-se que, desde o começo da escrita, ficou muito evidente que a linguagem oral, produzida na forma de uma cadeia de elementos sonoros, podia ser segmentada em unidades especiais, a que nós chamamos de *palavras*. No entanto, as poucas considerações acima já mostraram que tal unidade apresenta dificuldades em sua definição. Ela pode se referir a um simples monossílabo ou à composição de mais de um deles; pode remeter a um significado específico ou derivado, a partir da comparação com outros elementos ou por referência à maneira como as idéias são representadas na escrita, ou pode remeter a um processo de leitura cujo objetivo é apenas tornar a pronúncia transparente ou possível.

Com a normatização da escrita através do estabelecimento de uma ortografia, a palavra passou a ser a unidade principal da escrita, a responsável pela estruturação da categorização gráfica e funcional dos caracteres (Massini-Cagliari e Cagliari, 1999). Porém, quando é comparada com as unidades da linguagem oral, a noção de palavra

corresponde a várias unidades, de acordo com os modelos descritivos da Lingüística moderna.

A questão não é apenas semântica, mas fonológica e lexical. Embora nossa escrita ortográfica-alfabética use letras e não ideogramas, não é raro notar erros de escrita, sobretudo produzidos por crianças em fase de alfabetização, que revelam que o usuário segmenta palavras da linguagem oral atento a possíveis descobertas de significado, mesmo que bizarros (Massini-Cagliari, 1999: 33-39). Nesse caso, um significado detectado vale uma palavra, cuja ortografia logo se apresenta. Assim, uma criança pode achar que uma palavra como *visita* é composta dos significados de *vi* e de *Zita*. A criança não pensa que *visita* significa *vi Zita* na linguagem oral, mas acha que poderia escrever daquela forma. As escritas de *derrepente*, *com paixão*, em vez de *de repente*, *compaixão*, são exemplos que se encontram na prática escolar de alunos adiantados.

Certos erros que aparecem na alfabetização não aparecem na escrita de usuários já alfabetizados, mesmo que usem pouco a escrita. Por exemplo, uma pessoa alfabetizada não procura escrever a palavra *borboleta* em duas partes, mas, uma criança pode escrever *borbo leta*, porque acha que as palavras escritas não podem ser muito compridas. Além disso, os alfabetizados costumam juntar palavras, atribuindo a cada unidade de escrita um significado único, não passível de segmentação, segundo eles, guiando-se por unidades prosódicas maiores do que a palavra. Assim, escrevem *eraumavez*, *damãe*, *dotio*, *nacasa*, etc.

Apesar de a linguagem ser definida de maneira tão simples e clara, como a união de idéias e sons, significados e significantes, o reconhecimento das unidades geradas por essa definição está longe de ser simples e claro. A sintaxe cria um efeito cumulativo dessas unidades. A segmentação desses pedaços depende de muitos fatores. Os elementos prosódicos (entoação, grupo tonal, grupo clítico, etc.), a fonotática (a ordem dos segmentos na cadeia da fala), a estrutura silábica, o significado, a função sintática, etc. podem ser evocados para dizer onde começa ou termina uma unidade segmentada do contínuo da fala. Como vimos, para os usuários da língua, a esses fatores, junta-se, ainda, o conhecimento que têm da escrita e a maneira como lidam com as relações entre linguagem oral e linguagem escrita.

As considerações acima nos levam a concluir que segmentar a fala em unidades, sejam elas quais forem, não é tarefa fácil. Para um mesmo enunciado, há sempre várias possibilidades. A escolha de uma forma ou de outra é uma decisão que depende do que se quer ver e dos

instrumentais teóricos que os modelos descritivos da linguagem põem nas mãos do lingüista.

O desenvolvimento das teorias lingüísticas tem contribuído de maneira muito significativa para a compreensão da linguagem nos seus mais variados aspectos. Mas, o próprio fato de a história da Lingüística mostrar novas teorias, que surgiram em espaços de tempo muito curtos, às vezes, trazendo informações contraditórias com relação ao que elas queriam deixar de lado, próprias das abordagens anteriores, significa que não há muito consenso entre teorias e pesquisadores ou, melhor dizendo, não há, ainda, um conhecimento sólido e confiável nos resultados que se obtém aplicando essas teorias na descrição das línguas. A saída para certos impasses na interpretação de fatos que, de certo modo, comprometem o esquema teórico estabelecido, em geral, é feita através da invenção de modelos particulares para casos específicos ou, como costumam dizer os lingüistas, através de interpretações *ad hoc*.

O léxico é considerado o reino das exceções. A estrutura de uma língua comporta alguns componentes, como o sintático que cuida das relações entre as palavras e sua ordem nas frases, o semântico, que cuida da significação em todos os níveis (podendo-se incluir aí áreas da Lingüística como a Análise do Discurso), o fonológico que cuida dos sons e da maneira como eles materializam o significado através dos segmentos sonoros, dos elementos prosódicos, e o componente lexical, que nada mais é do que uma lista que contém os morfemas e as regras que permitem a combinação deles, formando palavras e outras unidades.

No léxico, encontram-se todos os morfemas da língua (menores unidades fônicas dotadas de significado). A relação do léxico com a sintaxe é mais ou menos fixa: uma palavra pertence a uma determinada categoria e com essa marca irá ocupar um lugar determinado na sintaxe. A relação do léxico com a semântica, além de apresentar valores literais e conotativos, caracteriza-se por grande mobilidade no tempo. Isso se vê facilmente comparando as descrições das entradas lexicais de dicionários antigos com modernos ou mesmo constatando que todas as palavras apresentam mais de um significado, o que mostra que a semântica é tipicamente cumulativa no léxico. A relação do léxico com a fonologia se define basicamente na formulação das regras fonológicas e morfofonológicas. Ao juntar morfemas para formar palavras derivadas ou compostas, há regras de reajuste fonológico, acrescentando elementos (epêntese), tirando (queda), ajustando (desvozeando, por exemplo), etc.

O componente lexical, em si, tem uma estrutura própria. Em primeiro lugar é uma lista de morfemas, cada qual com as marcas categoriais sintáticas, indicações semânticas e marcas fonológicas (para aplicação de regras específicas, em geral, representando exceções). Muito raramente, um item lexical sai do léxico para a fala sem sofrer modificações. Uma língua como o Português, por exemplo, junta morfemas para construir palavras, que irão, depois, aparecer na fala. Alguns morfemas trazem marcas gramaticais específicas, como as desinências verbais e nominais. Os verbos e os nomes, em Português, trazem sempre esses morfemas agregados. Há, ainda, morfemas que acrescentam significados especiais a outros morfemas, como os sufixos, acrescentando, por exemplo, a idéia de pequeno (diminutivos), grande (aumentativos), a idéia de pejorativo, etc.

Portanto, a palavra pode ser vista como uma unidade formada por morfemas, contendo um elemento básico, o radical, que carrega o significado principal e outros em acréscimo, os afixos, que definem sintaticamente a palavra ou a modificam semanticamente. Uma palavra pode também ser formada da união de duas ou mais palavras, como acontece no caso da composição vocabular.

Encontram-se também no léxico regras prosódicas, como a de atribuição de acento. Regras desse tipo têm sugerido uma estratificação do léxico em níveis. Além disso, pela maneira como funcionam no processo de derivação, algumas regras têm sido interpretadas como agindo ciclicamente, passando de um nível a outro na derivação. Uma regra cíclica é uma regra que se aplica em um nível, produzindo um certo resultado e, depois, continua se aplicando em outros níveis, acrescentando modificações sobre as modificações já obtidas anteriormente. Tem-se constatado também que algumas regras morfofonológicas precisam ser aplicadas antes de outras, caso contrário, perdem o contexto de aplicação e ficam inúteis, gerando formas indesejáveis de palavras. Por exemplo, algumas regras aplicam-se somente no final da construção das palavras, ajustando-as à fonética da língua. Por serem as últimas regras, são chamadas também de regras pós-lexicais. Essas regras não podem ser aplicadas antes, porque causariam uma mudança no contexto em momento inoportuno, impedindo a continuação do processo de derivação. As regras que se aplicam nos níveis iniciais costumam enfrentar o grave problema das exceções. As regras pós-lexicais são sempre absolutas e, portanto, não admitem exceção.

Tudo numa gramática pode sofrer variação e mudanças. Historicamente, alguns aspectos mostram-se mais ativos do que outros, acelerando mais a mudança. O léxico da língua, por ser um

conjunto aberto de elementos, costuma se mostrar muito vulnerável à variação e à mudança, quer em função do espaço, quer do tempo. O léxico incorpora facilmente elementos novos, oriundos de fatores externos (outras línguas) e internos (modificações nas regras do sistema). Ao fazer isso, estruturas que eram pouco representativas, (mas possíveis) na língua, podem passar a ter um volume grande de itens lexicais, criando embaraços para o sistema de regras de um determinado aspecto da língua. O sistema acentual do Português Arcaico ficou alterado, quando a língua incorporou palavras proparoxítonas, oriundas tardiamente do Latim, no Português Clássico (Massini-Cagliari, 1999a). O Português arcaico não tinha palavras proparoxítonas (a não ser, talvez, alguns nomes próprios, ainda com sabor latino, como Jerônimo, Cecília, etc.). Com as novas aquisições do léxico, a regra antiga passou a não ser mais completamente satisfatória e os novos dados tiraram o caráter geral da regra antiga, baseada no fato de as palavras serem paroxítonas e oxítonas.

O léxico é uma das principais portas, pela qual entram muitos elementos que vão constituir fatores de variação e de mudança de uma língua, re-organizando a ordem e a arquitetura do sistema. Em geral, tais elementos intrusos vão se infiltrando aos poucos e de muitas maneiras. As modificações não começam refazendo as regras, mas aceitando o novo ao lado do velho, no começo. Depois, o novo passa a ter o mesmo status do velho e as regras da língua precisam ser refeitas. As mudanças estão sempre dentro do possível, mas raramente dentro da lógica anterior do sistema. O sistema modifica-se para recuperar a situação de equilíbrio necessário, e oferece às novas gerações de falantes nativos um modelo novo.

A mudança lingüística faz com que o léxico seja um depósito que comporta elementos antigos e novos, o que, não raramente, traz consigo elementos contraditórios que a língua procura desfazer, mesmo que para isso leve um certo tempo ou tenha que apelar para modificações semânticas. Por exemplo, do Latim *cathedra* veio a palavra *cadeira*; mas, depois, voltou a aparecer a palavra *cátedra*, a qual trouxe mais um item lexical proparoxítono, porém, aceito com novo significado.

Como se nota, nos comentários apresentados acima, com o objetivo de começar a traçar um quadro geral dos fenômenos abordados no presente trabalho, a palavra é uma referência muito importante e central, seja de que lado se olhe o léxico e seus problemas. Para fins de descrição lingüística, a palavra cede lugar à noção de morfema. Assim, no começo da derivação, ou na forma subjacente da língua, vamos encontrar a noção de morfema e não de

palavra. Já na realização fonética, o que mais interessa é a noção de palavra e não de morfema.

O léxico, além de agregar a lista dos morfemas da língua, contém também as regras que se aplicam sobre os itens lexicais, transformando-os em palavras ou indicando a função sintática ou os atributos semânticos ligados diretamente a eles. Diferentes teorias apresentaram modelos específicos para o componente lexical, desde o modo como os dicionários se organizam até os modelos gerativos ou representativos, que apresentam uma arquitetura muito própria para o componente lexical.

Como vimos antes, as palavras têm estruturas facilmente identificáveis através de segmentação. As partes das palavras segmentadas, mantendo uma relação de significado e significante, são os morfemas. Morfemas podem variar no significado, na forma e nos elementos fonológicos que os compõem e criar alomorfas, ou seja, variantes morfológicas. Alguns casos de alomorfia criam interfaces com a sintaxe, com o léxico, e com a fonologia. A morfologia pode ser vista como concatenação de objetos (morfes) ou como regras ou processos (que atuam sobre estruturas sintáticas, semânticas e fonológicas).

2. Abstraindo da fala

A fala, como se apresenta aos interlocutores, representa um fato real físico, articulatório, perceptual, etc., que pode ser medido através de equipamentos adequados. A fala, como fenômeno lingüístico, traz consigo o sistema da língua em sua forma plena. O fenômeno lingüístico exige teorias poderosas para poder ser descrito adequadamente. Entretanto, entre o físico e o sistema existem os falantes e ouvintes, ou seja, o usuários desse processo. Os falantes não são os donos da linguagem, não podem modificá-la à vontade, uma vez que toda mudança da língua exige a aprovação de outros usuários e um certo tempo para a incorporação do novo ao sistema. Por outro lado, os usuários têm a capacidade de refletir sobre os mecanismos de produção, de percepção, sobre a própria estrutura da língua que falam e sobre a linguagem, em geral. Isso permite a eles falar sobre a linguagem e criar uma teoria a respeito de como ela é e de como funciona. Embora a tarefa de criar teorias lingüísticas pertença primariamente aos lingüistas, os usuários põem em prática certas ações que refletem indiretamente que tipo de teoria eles têm a respeito da linguagem.

Como vimos antes, um desses momentos importantes aconteceu quando o homem resolveu representar a fala através da escrita. Ao fazer isso, ele escolheu uma unidade da linguagem como base da escrita, que foi a palavra. Procedendo assim reconheceu que a escrita deveria ter como unidade central uma parte da fala que trouxesse consigo, de maneira facilmente segmentável, uma seqüência de sons à qual estivesse associado um significado claramente definido. Como a palavra podia se reduzir a um padrão prosódico bem definido, como um grupo tonal, por exemplo, e se relacionar facilmente com a referência indicada pela semântica, essa era a unidade ideal. Todavia, isso não foi conseguido sem resolver alguns problemas.

As palavras se relacionam umas com as outras. Tal relacionamento é tão grande que os lingüistas agrupam palavras chamadas “cognatas”, ou seja, “aparentadas” entre si. Se elas são aparentadas, precisam ter uma representação que lembre o leitor e o usuário da língua desse fato. Nesse sentido, os sistemas de escrita ideográficos guardam alguns traços comuns para certos conjunto de palavras. Os sistemas de escrita fonográficos, por sua vez, fazem isso a partir dos sons semelhantes que essas palavras têm. Tais fatos ficam mais claros, quando se observam palavras primitivas e derivadas ou compostas. Se existe uma forma simples, nada mais razoável do que usá-la na representação correspondente na forma composta. Por exemplo, na representação de uma palavra como *saleiro*, o reconhecimento da idéia e dos sons da palavra primitiva *sal* é um fator condicionante para que *sal* e *saleiro* sejam representados de maneira semelhante, no que têm de semelhante, quer nos sistemas ideográficos, quer nos sistemas fonográficos.¹

Em uma variedade da Língua Portuguesa, como a europeia, na qual os falantes usam a lateral alveolar na coda, como em [sal], a transposição da escrita ortográfica para a palavra *saleiro* não traz dificuldades, basta acrescentar o sufixo *-eiro* à palavra *sal*. Em uma variedade, como a brasileira, em que se diz [sau], a relação ortográfica entre *sal* e *saleiro* não é tão evidente para o aprendiz alfabetizando. De fato, as crianças escrevem *sau* e *saleiro* sem relacionar um item lexical com o outro.² Entretanto, uma pequena

¹ Obviamente, tal afirmação não representa uma regra geral, apenas mostra um fato comum, mas não absoluto, sobretudo nos sistemas ideográficos, em que predominam os aspectos semânticos.

² Em sua dissertação de mestrado, Santos (2000) mostra que as crianças escrevem a partir de uma relação que fazem de cada som com as respectivas

reflexão faz com que elas, baseando-se nos aspectos semânticos e fonológicos, reconheçam que as palavras *sal* e *saleiro* se relacionam. Os ortógrafos sempre procuraram essa relação entre as palavras com o objetivo de tornar a escrita mais ‘sistemática’.

Na busca de uma ortografia ideal, muitos fatores entram em jogo, como a etimologia, o parentesco entre as palavras, a natureza dos caracteres, etc. No século XVIII, quando a ortografia da Língua Portuguesa passou a representar a origem grega de algumas palavras, como *chimica*, *pharmacia*, *these*, etc., a representação ortográfica com os dígrafos CH, PH, TH fazia com que a escrita lembrasse que aquelas palavras tinham uma origem grega e que naquela língua, onde ocorria um CH havia uma letra “qui”, onde se escrevia com PH, havia uma letra grega ‘fi’, e assim por diante. Ou seja, o ortógrafo queria que os usuários da escrita fizessem associações entre a Língua Portuguesa e a Língua Grega. Podemos dizer de um modo grosseiro que, ao proceder assim, a ortografia mudou “sua forma de base”, recebendo uma informação extra, que não tinha, e agrupando as palavras de um modo diferente do que se fazia antes. As palavras gregas passaram a ter uma marca especial, em alguns casos.³

Analisando a ortografia atual da Língua Portuguesa, encontramos palavras do tipo apresentado a seguir:

belo	beleza
cola	colamos
pode	podemos

Nota-se que ocorre uma variação entre vogais [ɛ] e [e], em palavras como *belo* e *beleza*. No entanto, a ortografia optou por escrever a palavra primitiva e a derivada com uma única letra E que, na leitura, ficou com dois sons diferentes. Fato semelhante encontra-se nas palavras *cola* e *colamos*, em que aparece a alternância entre [ɔ] e [o]. Na palavra *pode* ocorrem as vogais [ɔ] e [i]. Porém, na palavra *podemos*, a escrita mantém as mesmas letras, apesar de a letra O

letras que acham que devem usar para representá-los e não é comum deduzirem a escrita de palavras simples a partir das derivadas ou vice-versa.

³ É curioso notar que os filólogos passaram a chamar essas representações ortográficas de *período etimológico*. No entanto, uma língua como a Portuguesa, que veio do Latim, tem na sua representação ortográfica, através das transformações ocorridas, a “etimologia” latina. A palavra *mesa* se escreve assim porque, em Latim, se escrevia *mensa*. Nesse sentido, uma grafia como *mesa* é tão etimológica quanto uma grafia do tipo *geographia*.

representar [o] ou [u] e a letra E representar [e] e não [i]. Como essas palavras têm uma relação muito forte entre si, a opção por uma única forma de representação ortográfica mostra uma escolha por uma “forma de base” comum. O processo de leitura se encarregará, depois, de achar a realização fonética adequada.

Continuando a observar a ortografia da Língua Portuguesa, encontramos exemplos como os abaixo:

sobe	subida
come	comida

Nesses casos, a ortografia diferenciou a palavra *subida*, mas podia perfeitamente escrever *sobida* (ou, então, escrever *cumida*).

Às vezes, aparecem algumas marcas para indicar diferenças, como se pode ver abaixo:

ótimo	otimizar
étimo	etimologia

Agora, as letras O e E receberam um acento agudo para sinalizar não apenas que a palavra é proparoxítona (que implica em uma marca de acento gráfico), mas também que se trata de uma vogal [ɔ] e [ɛ]. Podemos dizer que a “forma de base” continuou a mesma, apenas com o acréscimo de pequenas marcas (os ‘acentos gráficos’). Essas formas se diferenciam de palavras como

cômodo
ônibus

nas quais o acento circunflexo não interfere na pronúncia da letra O, apenas sinaliza a localização do acento na palavra. No caso anterior, o acento agudo representava duas funções no sistema ortográfico. Neste caso, o acento circunflexo representa apenas uma.

Vejamos, agora, os seguintes casos:

trouxe	trazer	trago
disse	dizer	digo

Existe um paralelismo entre essas formas verbais. No entanto, *trouxe* [trouzi] se escreve com X e *disse* [disi] se escreve com SS.⁴

Ao refletir sobre como se devem escrever as palavras, as crianças em fase de alfabetização acham que a letra E representa os sons de [e,ɛ,i] e que a letra O representa os sons de [o,ɔ,u]. Percebem também que existem os sons de [a] e de [ɐ̃], como em *gata* e *cama*. Acreditam que o som de [ɐ̃] deveria fazer parte das representações da letra E, pela semelhança com os outros (som não labializado, não posterior) e, por essa razão, costumam escrever esse som [ɐ̃] com a letra E: *cema* (cama), *banena* (banana), etc.⁵ Embora possa parecer um tanto estranho para os usuários da escrita da língua, esse modo de grafar a vogal nasal representa uma hipótese das crianças sobre uma certa “forma de base” da escrita de palavras na ortografia. Mais adiante, vão descobrir que a vogal [ɐ̃] pode ser escrita também com AN e passam, então, a escrever *canma* (cama). Essa é outra hipótese. Finalmente, descobrem que basta deixar o A diante de letra que representa a consoante nasal: *cama*. As várias opções de escrita representam hipóteses sobre como as palavras devem ser representadas na ortografia. Com a representação morfofonológica das formas de base dos itens lexicais ocorre algo semelhante e a analogia com o que acontece com a ortografia parece ilustrar bem o que se entende por uma ‘forma de base’ (no caso, aplicado ao sistema ortográfico da Língua Portuguesa).

Um exemplo interessante que, infelizmente, requer um pouco de pesquisa por parte de alguns leitores, é a história de como foi estabelecido, no século XVI, o sistema de escrita coreano moderno, chamado Han’gŭl.⁶ Há dois tipos básicos de representação gráfica: o traço e o ponto - um representando a terra e o outro o céu; um representando o *ying* e o outro o *yang*. Da união dos dois são formados os caracteres, de tal modo que o sistema represente uma ‘base comum’ com relação à qual se marca um traço distintivo por comparação com os demais caracteres, dando, desse modo, uma representação fonética precisa para as letras. Estas, entretanto, são

⁴ Na *Epítome da Gramática Portuguesa* que acompanha o Dicionário de A. de Moraes, o verbo *disse* vem escrito *dixe*. Tal forma já era encontrada no Português Arcaico.

⁵ Veja a dissertação de mestrado de Chaves (2000). Talvez seja por essa razão que algumas pessoas escrevem *terrapienagem* em vez de *terraplanagem*.

⁶ Veja Sampson (1985: 120-144) e Sang-Soon et alii (1990).

juntadas em sílabas para dar a impressão visual de um ideograma chinês.⁷ Porém, a questão mais interessante naquele processo de escrita apareceu quando foi definida a escrita das palavras. Para neutralizar a variação dialetal e permitir uma escrita mais duradoura, sem precisar de mudanças, a grafia das palavras acabou sendo representada por uma forma “mais neutra”, “mais abstrata”, semelhante ao que faz a Fonologia gerativa, estabelecendo as formas de base.

Do exposto acima, pode-se concluir que uma “forma de base” é uma hipótese sobre a representação de palavras no léxico da língua. Essa hipótese depende dos parâmetros estabelecidos pelas teorias lingüísticas. Diferentemente do que acontece na fala, as formas de base são abstrações que procuram generalizações possíveis ou desejáveis para organizar as palavras em conjuntos de determinados tipos e que permitam uma relação entre elas, através de regras. A passagem dessas representações para a pronúncia, na fonética, depende de um processo de derivação, como acontece na Fonologia gerativa ou de leitura em uma determinada variedade dialetal da língua, como acontece com a ortografia.⁸

3. O que é morfologia

A morfologia é a parte da gramática que estuda as palavras. Um modo de estudar a morfologia vê as palavras como um conjunto de elementos ligados a funções sintáticas (função sintática de certas unidades de sentido); outro vê a palavra como unidade central, com relação a qual outras partes da gramática se relacionam. Como as palavras são formadas de fonemas e exercem sempre uma função sintática, a morfologia se localiza, na gramática, entre a fonologia e a sintaxe. Historicamente, os estudos privilegiaram ora o aspecto fonológico, ora o aspecto sintático. A semântica das palavras ficou confinada aos estudos lexicológicos e lexicográficos, ou seja, aos dicionários tradicionais.

Historicamente, a morfologia apresenta três abordagens distintas. Na abordagem categorial, a palavra é o centro e o fundamento da língua e da gramática. Na abordagem semântica, a

⁷ A Coreia sempre usou o sistema chinês de alguma forma.

⁸ Nos estudos lingüísticos, o que conta é a fala e o sistema da língua. A ortografia foi usada apenas para ilustrar e não deve ser usada na análise e interpretação lingüística dos dados.

unidade palavra dá lugar à unidade morfema, definida como as menores unidades de significado. Na abordagem distribucional, segmentos idênticos, segmentáveis da corrente da fala, formam paradigmas, ou seja, conjuntos de elementos com uma mesma função, significado ou valor.

Para quem acha que a palavra é o ponto de partida, a morfologia é o estudo da formação dessas unidades e de seus usos na sintaxe. A morfologia é, assim, a descrição dos morfemas e seus modelos de ocorrência dentro das palavras. Esta é a interpretação mais antiga das gramáticas e típica dos filólogos. A própria gramática tradicional apresenta duas grandes divisões: uma que cuida das palavras como unidades e outra que cuida da frase, ou seja, da combinação de palavras nos enunciados. Aronoff (1976: 10) considera que os morfemas são unidades que realmente interessam somente quando vistas com relação às palavras que formam, o que coloca a palavra como unidade central da morfologia.

Para os estruturalistas (Harris, 1946), as palavras são segmentáveis em morfemas. Portanto, as palavras passaram a não ter importância para os estudos lingüísticos, sendo os morfemas as unidades básicas. A sintaxe deixou de ser uma seqüência de palavras para ser uma seqüência de morfemas.

Um outro modo de ver a morfologia vem desde Bloomfield (1926), que definiu o morfema como a menor unidade dotada de significado, ou seja, as menores unidades de significado que a língua identifica. O problema dessa abordagem é saber o que, de fato, significa “as menores unidades de significado”. Alguns lingüistas propõem a busca por alguns ‘significados mais profundos, subjacentes’ ligando morfemas aparentemente ‘sem sentido’ (Bybee, 1985: 4; Booij, 1987). Noções como Negação, Passado, Plural existem dentro do sistema da língua (referência à língua) e têm um *status* diferente do significado básico das palavras (referência ao mundo). Do ponto de vista morfológico, os significados podem ser lexicais (fazem referência ao mundo) ou gramaticais (fazem referência ao sistema da língua).

A distinção entre tipos diferentes de significado nas unidades morfológicas levou à distinção entre: palavra e morfema, por um lado e entre morfema e morfes, por outro. O morfe é a representação fonológica das unidades morfológicas. Por exemplo, a idéia de plural é o significado gramatical de um morfema. O elemento fonológico /S/, que representa o plural em Português é um morfe, ou seja, uma realização do morfema de plural, é o elemento fonológico que atualiza na fala (elemento material, concreto) o morfema (noção abstrata). A

distinção entre morfema e morfe é paralela à distinção entre fonema e alofone. A noção fonológica de variante, a morfologia criou a noção de alomorfia. Essa noção veio, de certo modo, aliviar os problemas das interpretações que se baseavam mais pesadamente na noção semântica ou no aspecto formal dos morfemas, podendo colocar numa mesma classe formas divergentes, como os morfes [s, z, is, iz] do morfema de plural do Português.

Uma abordagem diferente é a apontada por Nida (1946: 79) que diz que os morfemas são reconhecíveis por “semelhanças parcialmente diferentes entre expressões” (“*different partial resemblances between expressions*”). Assim, reconhece-se o morfema de passado nas formas verbais do tipo *comprado, ferido, vendido*, etc. ou o morfema *feliz* em *infeliz, felizmente, felizardo*, etc. Esse tipo de abordagem também apresenta alguns problemas de reconhecimento das unidades como ocorre em casos como o *in* de *infeliz* que não é o mesmo que o *im* de *importante*, ou de *inferno, interior*, etc. Palavras como *desabar, descobrir, descascar* apresentam um *des* cujo significado não é de um único e mesmo morfema. Em *desobedecer, desrespeitar*, o *des-* refere-se a um morfema ‘negativo’, mas em *descascar* já não se pode dizer o mesmo (cf., *desmoronar, destelhar, despedir*).

O critério distribucional para detecção de morfemas também é imperfeito e induz a erros: *religião* e *religar* não têm um mesmo morfema *re*. A decisão semântica nem sempre tem critérios claros para usar. Por exemplo, certos formativos lexicais apresentam significações diferentes: *portão* não é o aumentativo de *porta*, *lençol* não é diminutivo de *lenço*. Por outro lado, *urbano* e *cidadino* tem o mesmo significado e se referem a um único morfema. O grande problema dessas abordagens é como conciliar ‘forma’ e ‘conteúdo’ - que é, na verdade, o objetivo maior da morfologia. A morfologia precisa estabelecer um certo compromisso entre forma e conteúdo para poder ter uma unidade de trabalho que funcione. Somente pela forma ou pelo conteúdo não se consegue fazer uma morfologia que descreva a língua com precisão e de maneira adequada.

Uma definição pragmática diz que o morfema é o que os falantes costumam achar que constitui uma unidade simples e única, podendo aparecer em outras formações morfológicas. Aqui também encontram-se problemas de segmentação. Em Inglês: *cranberry, loganberry, strawberry*, etc. têm um elemento distinto *-berry* - tipo de fruta? e o que precede é também morfema? Certamente, se pode dizer isso de *straw* - mas não de *cran*, cujo significado, se é que existe, é desconhecido por todos os falantes.

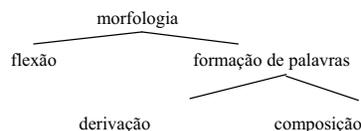
Para complicar o quadro, há ainda a interpretação etimológica dos morfemas: o que para um falante de hoje parece ser um morfema simples, um etimólogo pode achar dois morfemas formativos da palavra na história da língua (cf. *embora* e *em* + *boa* + *hora*; *você* e *vossa mercê*), etc. Para os estudos morfológicos, é preciso separar a abordagem sincrônica da diacrônica, caso contrário, fica impossível descrever o fenômeno com propriedade e precisão.

As noções de palavra, morfê e morfema podem ser exemplificadas na análise de uma palavra como *infeliz*, apresentada a seguir:

palavra	morfês	morfemas
<i>infeliz</i>	[ĩ - felis]	NEGATIVO 'FELIZ' ADJETIVO SINGULAR

Uma outra questão relevante para os estudos morfológicos é a formação de palavras. Essa questão é tão importante que a noção de produtividade é também um dos critérios para se definir o que é um morfema: morfema é o menor elemento (unidade) que pode aparecer em outras formações morfológicas.

Na formação de palavras, os morfemas apresentam três processos: flexão, derivação e composição. O primeiro cuida das desinências das palavras que se relacionam com as noções gramaticais de casos, gênero, número, pessoa, etc. O segundo trata dos afixos, ou seja, dos elementos que se agregam a um morfema de significado básico, modificando-o semanticamente em algum aspecto e possibilitando a mudança de categoria gramatical. O terceiro junta palavras já formadas para criar uma nova palavra, de significado diferente da combinação dos significados das palavras formadoras, embora possa haver alguma relação semântica entre elas (cf. *guarda-chuva*).



Bauer (1983: 34), atribui a noção de flexão à morfologia diretamente, não sendo um processo específico da formação de palavras.

No processo de formação de palavras, as categorias de flexão são mais constantes no significado e são altamente produtivas: todas as palavras de determinada categoria recebem determinadas flexões. O mesmo não acontece com o processo de derivação: nem todos os elementos de determinada categoria aceitam determinados afixos (cf. *barro* - *barrento*; *terra* - **terrento*)⁹.

Os compostos, por sua vez, tem peculiaridades nas línguas. Por exemplo, em Inglês, os compostos têm só um acento principal e admitem flexão apenas no último elemento: *bathroom*, **bath old room*, *old bathroom*, **baths room*, *bathrooms*. Podem ter elementos de qualquer classe e as formações nome + nome são as mais produtivas.

Há três abordagens morfológicas tradicionais: Palavra e Paradigma (*Word and Paradigm*); Item e Processo (*Item and Process*); e Item e Arranjo (*Item and Arrangement*). Palavra e Paradigma é a abordagem da gramática tradicional. A palavra é a unidade básica, central e fundamental da gramática. Metade da gramática refere-se à formação de palavras e a outra metade à sintaxe (juntar palavras).

A abordagem Item e Paradigma relaciona um item com um processo: a morfologia é, então, um processo (por exemplo, a idéia de plural). O morfema assume a posição principal na descrição. Todo morfema tem uma forma subjacente à qual são aplicados os processos. Às vezes, essa forma subjacente é o alomorfe de maior ocorrência na distribuição (Latim: *rex*, *regis*, *regi*, *regem*... [ks] aparece só no nominativo; nos demais casos aparece [reg] e 'reg' seria a forma subjacente). A gramática gerativa acrescentou processos aos itens lexicais (elementos do léxico) para poder derivar formas: *goose* + *PLURAL* = *geese*.

Item e Arranjo é a abordagem que considera a palavra uma seqüência de morfemas que podem ser segmentados. Aqui também, o morfema é a unidade fundamental. A respeito dessas abordagens, McCarthy (1991: 323) diz:

Palavra - Paradigma, Item - Processo e Item - Arranjo têm diferentes domínios de utilidade e nenhum modelo serve para todas as necessidades. Todos os três deixam certas áreas não resolvidas e as melhores características de cada um são, sem dúvida, essenciais em qualquer descrição lingüística completa.

⁹ Palavras precedidas por asterisco (*) são agramaticais, isto é, inexistentes.

Uma abordagem diferente da morfologia veio com a gramática gerativa (transformacional). Chomsky (1957: 32) interpreta a sintaxe como uma seqüência de morfemas, mas o centro da gramática ficou com a sintaxe e a fonologia. Lees (1960) deriva *manservant* de *The servant is a man*, usando uma abordagem transformacional da morfologia. Chomsky (1970) preferiu uma visão lexicalista, colocando no léxico os problemas de derivação e composição. O léxico é um conjunto de morfemas e das regras de formação de palavras (e as propriedades idiossincráticas dos morfemas e das palavras)

As regras de formação de palavras e as regras fonológicas são relacionadas em diferentes níveis do léxico ou *strata*. A cada nível aplicam-se as regras morfológicas e depois as fonológicas correspondentes ao nível. Ao final da derivação, passando por todos os níveis ou *strata*, todas as regras morfológicas e fonológicas foram aplicadas. Nesse modelo, o léxico contém as regras morfológicas e fonológicas (Pulleyblank, 1986; Katamba, 1989).

SEGUNDA PARTE

Noções Básicas de Morfologia

1. Forma e conteúdo

Apresentam-se, neste capítulo, algumas noções básicas de morfologia como forma complementar às informações trazidas anteriormente, na introdução deste livro. Os conteúdos escolhidos dizem respeito especificamente a definições de termos, tiradas da tradição da Linguística Moderna, cobrindo teorias tão diferentes como o estruturalismo e a fonologia lexical.¹⁰

O objetivo desta apresentação é dar subsídios para o leitor menos familiarizado com os estudos de morfologia. É muito importante ter diante de si essas noções e o quadro geral da morfologia aqui delineado ao passar para o estudo das outras partes.

O fato de haver propostas muito díspares com relação aos estudos de morfologia, às vezes, dentro de uma mesma vertente teórica, como o gerativismo, por exemplo, mostra que os linguistas não têm dedicado à morfologia o mesmo empenho que se tem observado em outras áreas, como a sintaxe, a fonologia e a semântica.

No final do livro, há muitas referências bibliográficas para consulta. Para tornar a leitura mais simples e fácil, as referências de apoio a noções gerais foram omitidas no texto, constando apenas da relação de obras apresentadas na bibliografia.

A definição mais precisa e técnica a respeito da linguagem diz que ela é constituída de *forma e conteúdo*. A forma é a realidade sonora da língua e o conteúdo são as idéias associadas a essas formas sonoras. Essa união pode ser encontrada em enunciados de todos os tipos e tamanhos. Uma palavra tem uma forma sonora e associada a ela uma ou mais idéias. Por exemplo, quando se diz *papéis*, o falante sabe que os sons dessa palavra remetem a idéia de *papel* e à idéia de mais de um, ou seja, à idéia de plural.

A noção de forma e conteúdo é central na morfologia. Seu objeto de estudo são as unidades menores de forma e conteúdo da

¹⁰ Recomenda-se a leitura de Lyons (1979, sobretudo das p. 139-218) para as pessoas não familiarizadas com alguns conceitos comuns da Linguística moderna relativos à morfologia.

linguagem, chamadas morfemas, os quais irão constituir as palavras da língua.

Os lingüistas costumam analisar a linguagem, vendo o conteúdo ou a forma, sempre um com referência ao outro. Assim, às vezes, o ponto de partida é a forma, os aspectos sonoros da língua. Para que eles tenham um valor lingüístico e não meramente físico, material, precisam estar a serviço da transmissão de alguma idéia ou significado. A simples constatação física das unidades e propriedades sonoras da linguagem não é suficiente para produzir uma análise lingüística. A função que os sons desempenham é que é relevante. A análise a partir da realidade sonora representa um começo fácil, pois trata-se de elementos facilmente controláveis, pela sua própria natureza. Nesse sentido, o trabalho do lingüista se volta para a comparação de fatos semelhantes e sua posterior interpretação, tendo em vista os resultados semânticos e sintáticos que decorrem deles.

O significado, por outro lado, é um elemento de fácil compreensão e controle por parte do falante nativo, mas que nem sempre é fácil de ser descrito ou mesmo de ser revelado. Dada sua natureza mental, o significado não tem partes e pode vir de maneira composta com outros significados que, às vezes, a língua permite separar, às vezes, não. Quando se diz *cardume* entende-se um conjunto de muitos peixes, mas a idéia de plural vem do conhecimento do mundo e não da linguagem. Por isso, *cardume* é uma palavra no 'singular', tendo como plural a forma *cardumes*, remetendo a uma realidade física diferente. Quando se diz *macacão* referindo-se a um tipo de roupa, a idéia de aumentativo já não faz sentido. Porém, quando a mesma palavra refere-se a um animal, chamado *macaco*, a idéia de tamanho grande vem logo associada. A palavra *calças* e *óculos* referem-se a objetos para os quais não associamos a idéia de plural, necessariamente. Todavia, do ponto de vista da estrutura da língua, essas palavras apresentam uma realidade lingüística, na qual há a marca de plural. A palavra *calças* passou a ser usada também na forma *calça*, mas a palavra *óculos* não gerou um singular lingüístico **óculo*¹¹ (cf., entretanto, a forma singular do artigo em *o óculos*).

O que aparece à primeira vista (ou na primeira audição) nem sempre acaba se revelando a verdadeira interpretação dos fatos da língua. As teorias contribuem muito para um certo afastamento que acontece entre a intuição primeira e a descrição formal final. Quando se diz *o banana* (pessoa 'preguiçosa, molenga'), *duzentos gramas*, *o*

¹¹ O asterisco diante de um enunciado representa uma forma agramatical, isto é, uma ocorrência inexistente na língua.

sofá, o avô, a avó, o pente, a diabetes, a dentista, o dentista, etc., nota-se que a idéia primeira a respeito da marca de gênero na língua é mais complicada do que a regra que diz que os nomes¹² terminados em *-a* formam o feminino e os terminados em *-o* formam o masculino. Por outro lado, nota-se que o artigo tem uma importância maior na determinação do gênero do que a terminação dos nomes.

A análise morfológica, como qualquer análise lingüística não pode perder de vista o fato de que a língua tem uma estrutura e que toda interpretação dessa estrutura, em qualquer nível ou parte, deve levar em conta sempre a forma e o conteúdo como dois lados indissolúveis de um mesmo objeto. Quando se privilegia a forma ou o conteúdo é pura abstração dessa realidade, para fins metodológicos. No final, os dois caminhos devem se unir novamente.

2. Segmentação

Segmentação é o processo lingüístico pelo qual um enunciado é decomposto em partes ou unidades menores. Partindo de uma frase, podemos encontrar duas partes sintáticas: o sujeito e o predicado. Isso é uma forma de segmentação. Assim, na frase *a pedra caiu*, pode-se segmentar esse enunciado em 'sujeito': *a pedra* e em predicado: *caiu*. Semanticamente essas partes podem ser interpretadas como 'tópico' e 'comentário'. Na frase acima, o sujeito gramatical é *a pedra*; mas, se quisermos analisar o enunciado do ponto de vista da 'lógica' das idéias, vão concluir que *a pedra* é o paciente de uma ação causada por um sujeito lógico que não foi revelado no enunciado acima. Esse sujeito só existe na concepção lógica do que foi dito. É um lugar vazio ou não preenchido ou especificado naquela frase. É possível segmentar um enunciado de diferentes pontos de vista e, ao fazer isso, é possível encontrar alguns 'buraco' ou lugares não preenchidos. Se a frase fosse *a pedra caiu com a chuva*, a expressão *com a chuva* tem uma interpretação diferente, se vista com os olhos da sintaxe ou com os da lógica. Toda segmentação, entretanto, mantém intimamente ligado dois aspectos da linguagem que são inseparáveis: a forma e o conteúdo. Como os enunciados representam uma sucessão de forma com seus respectivos conteúdos (ou vice-versa), é possível destacar cada parte que tem essa dupla face, constituída de forma e conteúdo. Esse procedimento revela que a linguagem se estratifica em diferentes

¹² O termo 'nome' na Lingüística significa *substantivos e adjetivos*.

níveis, de diferentes extensões e de diferentes funções. Além disso, podem ser vistos de diferentes ângulos ou pontos de vista.

Um enunciado traz consigo muitas informações de vários níveis, desde a menor unidade constituída de forma e conteúdo, como os morfemas, até as estruturas sintáticas das frases, e os elementos que fazem interface com tudo o mais que a linguagem carrega, como a pragmática, a sociolinguística, a lógica, a semântica argumentativa, o discurso, a estilística, a literatura, etc. Em todo estudo que envolve a linguagem, há de considerar sempre um aspecto que é a forma e outro que é o conteúdo, sendo ambos dois lados de uma mesma moeda e de relação necessária para os sistemas linguísticos. Sem essa relação, as línguas não se agüentariam e se desfariam na incerteza e na falta de suporte adequado para ser transmitida e recebida pelos falantes. Os lingüistas estão, pois, segmentando a todo instante, para poder definir um ponto de vista e as unidades com as quais querem trabalhar. A frase é a unidade mais transparente para os falantes. Unidades maiores, como as do discurso, e unidades menores, como os morfemas apresentam dificuldades e exigem cuidado especial na sua manipulação.

Nesta obra, a segmentação dos enunciados em morfemas é uma questão básica e presente a todo instante. Pode-se, por exemplo, isolar de um enunciado do tipo *preciso de quatro lápis de cores diferentes*, a palavra *lápis*. Qualquer outra segmentação nesse pedaço, destrói a exigência de se ter ao mesmo tempo forma e conteúdo. O *lá*-de *lápis* não significa nada, o *-pi-* também não, e o *-s*, neste caso, não representa o plural, como acontece comumente na Língua Portuguesa. Com relação à palavra *quatro* também não se pode fazer nenhuma segmentação morfológica. Nesse caso, o *-o* final não se refere ao gênero masculino e, embora a palavra represente mais de um, não existe uma marca formal de plural, como se encontra, por exemplo, no *-s* de *casas* ou no *-es* de *rapazes*. Por sua vez, a palavra *rapazes* pode ser segmentada em duas unidades: *rapaz* e *-ez*. A primeira parte representa um ser e a segunda uma função gramatical de marcação de plural. Na palavra *rapazes*, não encontramos nenhum elemento da forma (fonemas) ao qual possamos atribuir a função de gênero, que os substantivos da língua têm. Para esse significado gramatical existe um vazio da forma. Embora apresente algumas dificuldades, podemos reconhecer na palavra *casas* um segmento no elemento da forma *-a-*, na segunda ocorrência, como um marcador de gênero feminino, função que o primeiro *-a-* claramente não tem. Dessa forma, podemos concluir que a palavra *casas* tem três elementos da forma associados a três significados: *cas + a + s*.

Como as línguas se utilizam de um inventário relativamente pequeno de unidades básicas, formadas pela união de forma e conteúdo, para produzir uma quantidade muito grande de outras unidades que agregam também formas e conteúdos, o que é segmentável em um caso pode não ser em outro, por causa, em geral, do significado e, às vezes, da forma também. Quando se diz *casa de saúde*, referindo-se a um 'hospital', por causa do significado singular que o enunciado carrega, não se pode segmentar a expressão em partes menores sem destruir a unidade de forma e conteúdo que a expressão representa. No entanto, para efeito de análise, assim como se diz que a forma de uma unidade morfológica é constituída de fonemas e seus alofones de traços distintivos, também se pode dizer que a palavra composta *casa de saúde* é composta das palavras *casa + de + saúde* e que o elemento *casa* é formado de *cas + a*. A diferença entre a expressão *casa de saúde*, que é uma palavra composta, e a expressão *casa de Pedro* reside no fato de a primeira manter sempre os mesmos elementos representando uma idéia que, às vezes, pode ser até comparada com outra expressão não composta, como *hospital*. No segundo caso, não ocorre uma união semântica tão forte entre os elementos e, por essa razão, a língua não dispõe de uma expressão simples para traduzir aquelas idéias em uma única palavra. Expressões compostas podem ser formadas da soma de vários elementos, gerando um todo único, como, por exemplo 'o *toma lá e dá cá* do Governo'. A Língua Portuguesa usa o artifício de colocar hífen na escrita para representar esse tipo de formação morfológica: 'O *toma-lá-dá-cá* do Governo'. Uma evidência da formação de um todo único com esses elementos se vê na função sintática que essas expressões desempenham. No exemplo acima, a expressão *toma-lá-da-cá* vale por um substantivo e, portanto, pode receber o artigo ou outro determinante, ocupar a função de sujeito, etc.

Quando se comparam palavras da língua, torna-se menos difícil fazer as segmentações e localizar as unidades morfológicas. Veja os seguintes vocábulos: *por*, *dispor*, *indispor*, *disponível*, *indisponibilidade*. Nota-se que há duas formas para um mesmo significado: *po* e *pon*, assim como *ivel* e *ibil*.

Segmentar enunciados em frases e estas em palavras é uma tarefa que todo falante nativo faz com certa facilidade. A segmentação em unidades menores do que as palavras traz algumas dificuldades. Ao segmentar uma palavra procuramos antes de mais nada o seu núcleo, através de seu significado principal. Este, por sua vez, vem modificado por outras unidades que lhe alteram o significado em algum sentido ou lhe acrescentam informações gramaticais

importantes para o uso da palavra na sintaxe e no discurso. Nem sempre a idéia principal repousa em uma unidade sonora que pode ser dita isoladamente e isso dificulta a segmentação e a análise. De acordo com a profundidade no tempo dos conhecimentos lingüísticos a respeito de uma determinada língua, a interpretação pode chegar a graus sofisticados de análise. Em princípio, deveria bastar a intuição do falante nativo. E na maioria das vezes isso funciona bem. No entanto o apelo a fatos históricos pode enriquecer e esclarecer muito a realidade da língua no momento atual. Além disso, existe uma certa lógica na interpretação das partes de uma palavra. Por exemplo, as unidades que denotam uma negativa precisam definir a que parte ou partes da palavra elas se referem. A título de ilustração, vamos segmentar a palavra *indecomponibilidade*, que tem a ver com a idéia de 'por' associada à idéia de 'junto', lembrando a palavra *compor*. Porém, juntando a unidade *-de-*, que denota uma negativa, tem-se a palavra *decompor*, isto é, não 'por junto', ou seja, 'separar'. Para mudar a categoria de verbo para adjetivo, foi acrescentada a unidade *-vel-*, gerando a forma *decomponível*. Nota-se que a forma sonora do verbo 'por' ficou alterada: *poni*. Poder-se-ia interpretar também a forma adjetiva como sendo *decompon* + *ível*. Essa vogal, na verdade, aparece como um tipo de vogal temática que define as conjugações verbais. O adjetivo acrescentou uma idéia a mais ao significado original, a de que se trata de algo 'susceptível de ser decomposto'. Para transformar esse conjunto de idéias em uma propriedade, a forma morfológica anterior passa a ser um substantivo, com o acréscimo da unidade *-idade*. Com o acréscimo do significado de propriedade e da mudança de categoria, termos, agora, uma palavra como *decomponibilidade*. Nota-se, mais uma vez, que houve uma mudança em alguns elementos sonoros, transformando a unidade *-ível* em *-ibil*. Essa mudança sonora, como a anterior, não alterou o significado desejado. Por fim, a esse conjunto de idéias, formado pela "propriedade de algo susceptível de não ser posto junto" (ou seja, "ser separado"), foi acrescentada mais uma vez a idéia de negação que se refere de modo especial à idéia de 'propriedade', significando, portanto, a 'não propriedade... etc.' Como foi visto, a primeira negativa tinha como objetivo negar a relação que tinha sido estabelecida entre a unidade vizinha à direita e a seguinte, na mesma direção. A outra negativa tinha como objetivo negar a informação da unidade agregada mais à direita. Se a unidade negativa do tipo *de(s)* viesse antes da unidade *in*, a negativa mais à esquerda teria como objetivo negar a que lhe viesse logo à direita, como se pode notar em uma palavra como *desincompatibilização* que significa negar a

negativa formada pela unidade *in* em *incompatibilização*. A análise não se esgota nos comentários acima. Na formação da palavra anterior, passamos por várias categorias de palavras e como as categorias tem estruturas próprias de flexão, tal fato podia ter sido levado em consideração na análise. Ainda mais, as relações entre as unidades morfológicas na constituição das palavras, além do aspecto semântico 'lógico', costumam apresentar também restrições puramente lingüísticas. Por exemplo, poder-se-ia fazer uma hipótese para saber se há alguma relação entre as unidades que precedem o núcleo e as que seguem, por exemplo, se o fato de uma palavra terminar por *-idade* ou por *-ção* implica na escolha de determinados elementos ou em determinada ordem de elementos das unidades morfológicas precedentes, sobretudo se antecederem o núcleo e tiverem um significado de negação. Além da sintaxe da frase, que cuida da ordem das palavras nos enunciados, existe uma espécie de sintaxe da palavra, que se ocupa com a ordem das unidades morfológicas constitutivas das palavras. Na formação de palavras, não é raro aparecer mais de uma opção em determinados momentos. Por exemplo, da forma *compor* pode formar ou *decompor* (verbo) ou *componível* (adjetivo) - a forma *decomponível* vem por qual caminho? vem de *decompor* + *ível* ou de *componível* precedido de *de*? Embora o *in* negue o significado de 'propriedade', veiculado por *idade*, a palavra podia receber um ou outro antes, mostrando dois caminhos de formação.

3. A noção de morfologia

Uma pessoa comum consegue segmentar palavras pelo significado de alguns elementos, cujo reconhecimento é mais fácil ou familiar, comparando palavras entre si. Por exemplo, *algodoal*, *laranjal*, *bananal*, etc. revelam que há a referência a um produto agrícola: *algodão*, *laranja*, *banana* - e uma referência à plantação desses produtos. Este último significado aparece porque a palavra original foi acrescida de um sufixo *-al*. O significado é, pois, um bom ponto de partida para segmentar palavras e revelar sua estrutura interna. Porém, alguns tipos de palavras apresentam uma complexidade de elementos que não podem ser identificados apenas através da análise do significado, com a mesma facilidade com que se fez acima. Por exemplo, há as palavras *elétrico* e *eletricidade*. A palavra *elétrico* tem um sufixo *-ico*, formador de adjetivos (cf. *rítmico*, *prosaico*, etc.) e remetendo a uma idéia geral de 'qualidade'.

Esse significado pode ser encontrado também em todos os adjetivos qualificativos, mesmo que não terminem com o sufixo *-ico*, como *bela, comprido, alta*. A noção de 'qualidade' associada ao sufixo *-ico* é de natureza diferente daquela associada aos adjetivos qualificativos, que não têm o sufixo *-ico*? Por outro lado, a palavra *eletricidade* é derivada de *elétrico* e, em vez de ter uma oclusiva velar no sufixo, apresenta uma fricativa alveolar. Nas palavras *pão, padaria* e *panificadora*, há o significado de 'pão'. Porém, como segmentar essas palavras, associando sons com a idéia de pão? Isso nos leva a perguntar até que ponto é possível segmentar as palavras, guiando-se apenas pelo significado? A estrutura das palavras é formada por um complexo de informações semânticas, sintáticas e fonológicas, que é preciso distinguir, para que as unidades morfológicas sejam identificadas e caracterizadas adequadamente.

Ao segmentar uma palavra em unidades menores dotadas de significado, seja de qual natureza for, identificamos os morfemas dessa palavra. Uma palavra como *menina* é formada pelos morfemas *menin + a*. O primeiro refere-se à idéia de criança e o segundo à informação gramatical, segundo a qual essa palavra pertence à categoria feminino. Se a palavra fosse *menino*, teríamos um *-o* que é um morfema que identifica a palavra como sendo da categoria masculino. No caso da palavra *laranja* ainda é fácil de reconhecer que o *-a* final indica o feminino. O mesmo já não acontece com uma palavra como *dentista* numa frase como *Pedro é um dentista*. O termo morfema representa uma noção abstrata, a unidade morfológica com todos os seus ingredientes (significados, funções, categorias, usos, etc.). O termo morfê refere-se à seqüência de fonemas (ou alofones) que carrega o morfema, sendo, pois, um elemento ligado mais diretamente à fala do que às estruturas mais internas do sistema da língua. Nas palavras *elétrico* e *eletricidade* há um único morfema, realizado fonologicamente através de dois morfês: um *-ik-* e outro *-is-*.

Um morfema pode ser também uma palavra independente, como os números, em geral: *nove, dez*, etc. Na Língua Portuguesa, entretanto, a maioria dos morfemas lexicais não podem ser usados como palavras independentes (cf. **menin*).

Para quem considera que, em palavras como *estrutural*, há apenas um /L/ final que transforma um substantivo em adjetivo, tem-se apenas um morfema. O arqúfonema lateral, nesse caso, não seria um morfema, uma vez que ele não tem um significado próprio, mas apenas uma função morfológica. Para quem interpreta o sufixo como sendo /aL/ e não apenas /L/, como em *laranjaL*, na palavra *estrutural* há dois morfemas. Resta, entretanto, saber se o sufixo /aL/ de *laranjaL*

é exatamente o mesmo que se encontra na palavra *estrutural*, por causa do significado associado a essas palavras. Mas, esta é outra questão, que será vista em outro lugar. No presente trabalho, tal interpretação não é aceita. Portanto, o arqúfonema deste caso representa um morfema e a ele está associada a função de mudança de categoria morfológica da palavra *estrutural*.

Em resumo, a morfologia é o estudo das palavras. Toda língua tem palavras. Todo falante é capaz de segmentar o contínuo da fala de um enunciado em unidades menores, até atingir um ponto em que o que diz perde o sentido. A morfologia, pois, trata das relações entre forma e conteúdo, sons e idéias, definindo unidades. A palavra é a menor unidade desse tipo que pode ser falada isoladamente e reconhecida pelo falante como um fato perfeito de sua língua. Um enunciado como *Era uma vez um menino e uma menina* permite que o falante de Português reconheça oito palavras, cada qual podendo ser dita isoladamente, trazendo à fala uma seqüência de sons e um significado. O mesmo falante reconhece que há uma semelhança estrutural entre as palavras *um* e *uma*, ou entre *menino* e *menina*, uma vez que a diferença refere-se apenas ao fato de *um, menino* referirem-se ao masculino e *uma, menina* referirem-se ao feminino, categorias gramaticais que a língua usa. Embora o falante reconheça em uma palavra como *menina* que existe o significado de criança e de feminino, não faz sentido dizer apenas *menin* ou *a*, para representar o conceito de criança e o de feminino.

A morfologia tem a ver, então, com a formação das palavras da língua quer do ponto de vista da fala, quer do sentido ou da estrutura gramatical. Por essa razão, a morfologia está intimamente ligada à fonologia, à semântica e à sintaxe. É muito comum entre os lingüistas distinguir morfologia de léxico. A morfologia estuda os princípios que regem a formação das palavras, ao passo que o léxico é apenas uma lista. Estudos recentes, entretanto, como a Fonologia Lexical, têm procurado ver o léxico como uma lista organizada de um certo modo, cuja estrutura é importante, por exemplo, para a aplicação de regras fonológicas ou sintáticas.

O léxico representa uma lista de palavras. O estudo a respeito da elaboração de dicionários chama-se lexicologia, sendo a lexicografia a própria obra de elaboração de dicionários. Os dicionários, em geral, preocupam-se apenas com o significado das palavras e com um tipo de categorização sintática tradicional (substantivo, adjetivo, verbo, preposição, etc.). Outros tipos de dicionários são mais raros e especializados, como os dicionários de regência verbal e nominal. A regência é uma característica sintática (e

semântica) de algumas palavras que, como a indicação de categorias nominais e verbais, aparece dentro das palavras, sem poder ser segmentada, misturando-se com o significado básico da palavra. Assim, a palavra *menina* é um substantivo, mas tal fato não pode ser atribuído a um segmento da palavra, como acontecia com a noção de feminino. O fato de uma palavra como *necessidade* exigir um complemento precedido pela preposição *de* representa o mesmo tipo de fenômeno. Neste caso, porém, já não se pode dizer que o último segmento da palavra tenha como função indicar o feminino.

A morfologia, portanto, distingue-se do léxico, da fonologia, da sintaxe e da semântica. A noção de palavra, por sua vez, une todas essas áreas da gramática. A especialidade da morfologia é cuidar de como a palavra se forma. Não se trata de uma simples seqüência de fonemas, nem da materialização de uma idéia, representada pela seqüência de fonemas. Uma palavra apresenta uma sintaxe muito especial, delimitada na sua extensão e função das partes. Nesse sentido, a morfologia é também uma espécie de sintaxe ou uma fonotática mais sofisticada.

As relações entre palavras e semântica têm um interesse muito secundário no presente estudo. O que mais interessa são as relações (interfaces) entre a morfologia e a fonologia e entre a morfologia e a sintaxe. Dentro da abordagem da Fonologia Lexical, a morfologia atribui ao léxico uma importância que não tem em outras teorias. Tradicionalmente, a morfologia tem sido vista por uns como um conjunto de palavras com significados gramaticais ou não, que ocupam um determinado lugar nas frases. Por outros, tem sido vista como um conjunto de regras que definem e estabelecem essas funções, sem precisar fazer referência específica à maneira como isso ocorre fonologicamente. Por exemplo, em Inglês, o particípio passado termina em *-ed* em algumas palavras e em *-en* em outras (cf. *studied*, *broken*), formando dois conjuntos específicos (não contáveis / contáveis?). O fato de haver as realizações fonológicas diferentes é irrelevante para definir a função de particípio passado (ou da idéia verbal de 'passado'). Para esses lingüistas, a morfologia não passa de um conjunto de regras, enquanto para os anteriores, não passaria de uma lista de itens lexicais.

As relações entre a morfologia e a sintaxe têm sido vistas como se o componente morfológico contivesse apenas uma lista de itens, estando a formação de palavras a cargo da fonologia e da sintaxe, ou seja apenas como um léxico, ou como sendo até certo ponto sobrepostas. Uma outra maneira, mais radical, de encarar os fatos dentro da abordagem do segundo caso, considera que a

morfologia contém um léxico e um conjunto de regras. Os itens lexicais trazem as informações fonológicas, semânticas e sintáticas. Neste último caso, um item lexical traria consigo a informação de que categoria pertence (nome, verbo, etc.) e, em alguns casos, também as indicações sintáticas de regência e concordância. Sendo assim, a sintaxe nada mais acrescenta do que a ordem em que devem aparecer os itens lexicais, quando as relações entre eles já vêm expressas juntamente com as especificações do item lexical a que pertencem. Por exemplo, uma língua como o Latim, rica em declinações, contém morfemas que direcionam algumas relações sintáticas, definindo o que é sujeito, predicado e complementos. Por essa razão, a ordem das palavras na frase tem uma liberdade muito grande. Obviamente, o mesmo fenômeno poderia ser visto, como o é mais comumente, do ponto de vista da sintaxe: o componente sintático define uma estrutura de frase, a qual é preenchida com os itens lexicais correspondentes. Visto desta forma, os itens lexicais são apenas indexados sintaticamente, formando paradigmas para um uso adequado pela sintaxe. A Lingüística moderna deu tanta ênfase à sintaxe, sobretudo com os trabalhos de Chomsky e dos gerativistas, que uma visão morfológica das relações sintáticas chega a parecer muito estranha. No entanto, o fato de as línguas que dispõem de um sistema muito amplo e abrangente de flexão permitirem uma grande liberdade na ordem das palavras (na superfície), ao contrário das que apresentam sistemas muito limitados de flexão e pouca liberdade na ordem das palavras, favorece a possível interpretação de muitos fenômenos sintáticos como oriundos da própria morfologia. Certas palavras obrigam a presença de uma outra, precedida de preposição, como *necessidade* em *O homem tem necessidade de companhia*. A criação do sintagma preposicional está prevista na indexação sintática da palavra *necessidade*, ao contrário da palavra *homem* ou *companhia*, que não vêm indexadas com essa obrigação sintática. Se houver um sintagma preposicional após essas palavras, isso se deve a razões sintáticas e não morfológicas.

Enfim, por se relacionar com outros componentes de maneira muito profunda, a morfologia pode ser deixada de lado completamente ou até mesmo ocupar um lugar que, tradicionalmente tem sido atribuído aos outros componentes com os quais desempenha um papel de interface no sistema das línguas. Essa é uma discussão não resolvida e que surge, de vez em quando, a pretexto da análise de alguns fatos ou de alguns tipos de línguas.

4. A noção de palavra

Palavra é a unidade de escrita por excelência, mas não se ajusta às unidades lingüísticas de maneira perfeita. No entanto, é uma unidade usada comumente, sobretudo porque também os lingüistas se utilizam da escrita em seus trabalhos e, muitas vezes, a reflexão começa a partir de um enunciado escrito, embora o objetivo primordial da ciência lingüística seja a linguagem oral, sendo a escrita apenas uma representação da linguagem oral. Definida do ponto de vista da escrita, a palavra é a unidade de escrita separada por espaço em branco ou delimitada de outra forma. Trata-se de uma unidade segmentada para facilitar a leitura. Quando se compara essa unidade de escrita com as unidades de análise lingüística, constata-se que uma palavra corresponde, em geral, a um morfema lexical acrescentado de seus afixos (prefixos, sufixos e desinências flexionais). Visto de outro ângulo, uma palavra costuma corresponder a uma unidade prosódica que porta apenas um acento principal. Essas duas maneiras de definir a palavra, por comparação, apresentam falhas mais ou menos graves. As preposições, os artigos e os clíticos são morfemas gramaticais, não possuem acento principal, a não ser raramente, e, no entanto, formam palavras isoladas na escrita do Português. Como exemplo, observe o enunciado *O chefe te chamou de chato*. Por outro lado, palavras muito longas podem ter dois acentos principais foneticamente, ou seja, definir a constituição de dois pés rítmicos, como

//l ^ apli /caram a /indisponibili /dade dos di /reitos po /líticos //. ¹³

Palavras compostas representam uma unidade de forma e conteúdo, mas podem vir representadas por mais de uma palavra na escrita, como em *cirurgião-dentista*. Todavia, nem toda palavra composta vem separada na escrita, como em *malmequer*, *aguardente*, *vaivém*, *fotocópia*, *automóvel*, etc.

Estudos lingüísticos têm demonstrado que a noção de palavra é bastante intuitiva para os falantes nativos, mesmo analfabetos ou falantes de línguas ágrafas, os quais conseguem segmentar o contínuo da fala em unidades a que podemos chamar de palavras. A segmentação é feita, em geral, por um processo distribucional que permite, através da comutação de elementos de um mesmo paradigma, separar um elemento de outro, dando-lhes o *status* de unidade independente. Por exemplo, fazendo a comutação de elementos

¹³ O sistema de transcrição da entoação e do ritmo segue Cagliari (1982).

podemos constatar os seguintes enunciados: *o lápis preto quebrou*; *o lápis preto / sumiu*; *o lápis / azul / quebrou*; *a caneta / preta quebrou*; *aquele / lápis preto quebrou*, o que mostra que cada palavra pode ser substituída na frase.

Apesar das restrições acima, a noção de palavra, emprestada da escrita, como foi dito, é muito usada pelos lingüistas, mas requer sempre uma certa atenção e cuidado. Para alguns, existe uma distinção sutil entre *palavra* e *vocabulo*; para outros, os dois termos são sinônimos perfeitos. O termo *vocabulo* é usado, sobretudo na análise gramatical, referindo-se à palavra, sem levar em conta seu significado ou categoria. A palavra *termo* equivale a *vocabulo* nessa acepção. O termo *enunciado* refere-se a um conjunto de palavras articuladas sintaticamente, podendo formar uma frase ou uma unidade maior ou menor. O termo *expressão* ou *sintagma* refere-se a um grupo de palavras que têm uma certa relação entre si, como as palavras que integram uma categoria sintática do tipo sujeito, predicado e complemento..

5. A noção de morfema

Morfema é a menor unidade fonológica que carrega um significado lexical (do mundo) ou gramatical. A palavra *infelizes* traz três significados: NEGATIVO + ‘feliz’ + PLURAL. A esses significados podem ser atribuídos os sons representados na escrita pelas seguintes letras: *in* + *feliz* + *es*. O significado NEGATIVO é fruto de uma constatação que se faz como um ser inserido no mundo real. Por exemplo, em vez de *infeliz*, para denotar o estado de uma pessoa, poder-se-ia dizer também *triste*. O significado de *feliz* também tem uma referência direta ao conhecimento de mundo. Esses dois morfemas são, pois, de natureza lexical. A idéia de plural, embora também possa ter uma referência aos fatos do mundo real, por exemplo, pode-se dizer *dez pessoas*, é representada de maneira especial pela estrutura da linguagem, estabelecendo relações sintáticas entre palavras, o que não ocorre, por exemplo, com o *in*-NEGATIVO. O significado ‘gramatical’ aparece nos morfemas chamados gramaticais. Em suma, deve-se interpretar como *morfema* toda seqüência fonológica à qual se pode atribuir um significado lexical ou gramatical. Às vezes, um morfema pode ter mais de um significado que não podem ser atribuídos distintamente a partes de seus elementos fonológicos. No exemplo acima, a idéia de plural tem uma dupla função (ou significado): a de remeter ao mundo e à estrutura sintática

da linguagem. O mais comum, entretanto, é observar que um morfema tem mais de uma função ou significado gramatical. Por exemplo, nos verbos da Língua Portuguesa, as noções gramaticais de tempo e modo não podem ser segmentadas, estando presentes em uma unidade morfológica. Em *falava* e *falasse*, o morfema *-va* representa o imperfeito do indicativo e o morfema *-sse*, o imperfeito do subjuntivo. Não é possível separar a idéia de ‘imperfeito’ verbal da idéia de ‘indicativo’ ou de ‘subjuntivo’. Em outras palavras, não faz sentido dizer que *-v-* está associado à idéia de ‘indicativo’ e *-a*, à idéia de imperfeito - ou vice-versa. Mas, faz sentido dizer que a seqüência *-va-* representa o imperfeito do indicativo. Os morfemas cumulativos são sempre interpretados como exercendo duas funções gramaticais específicas.

Os morfemas lexicais são interpretados sempre como morfemas simples e não cumulativos. Aqui entra em jogo a discussão sobre a arbitrariedade do signo. As línguas fazem recortes semânticos de acordo com a maneira como interpretam o mundo. Como a ambigüidade é um dos fenômenos mais comuns e corriqueiros da linguagem (basta ver os dicionários), se a uma mesma seqüência de fonemas se puder atribuir mais de um significado, cada um deles constituirá um morfema à parte, um novo item lexical. A noção de morfema cumulativo apoiada na soma de significados para um mesmo enunciado não é contemplada pela Linguística. Por outro lado, isso poderia ser feito dentro de um quadro teórico específico e montado com essa finalidade. Os nossos dicionários costumam apresentar as entradas pelo significados diferentes de uma ‘mesma palavra’. O dicionário Chambers apresenta as entradas pelo aspecto fonológico semelhante dos itens lexicais.

O problema da ambigüidade reside no fato de o resultado final poder ser a convergência de coisas de origens diferentes. Assim, a palavra *casa* é ambígua fonologicamente falando, porque uma forma é substantivo ‘casa’ e outra é verbo ‘casa’ (de ‘casar’). Neste caso, a forma semelhante que eles têm na pronúncia é uma coincidência fonética. Esse tipo de palavra costuma, em geral, apresentar uma distribuição diferenciada, uma vez que, no exemplo acima, uma forma é substantivo e outra é verbo, duas categorias que exercem funções sintáticas que não se misturam. Um exemplo de natureza diferente é a ambigüidade que se nota na palavra *casa* em expressões do tipo: ‘a casa de João fica na Rua 12’ e ‘a casa do botão rasgou’. As duas ocorrências da palavra *casa* remetem a significados diferentes, porém, mantendo a mesma categoria de substantivo. Portanto, podem preencher a mesma função sintática. A ambigüidade que se nota, neste

último caso, é diferente - em certos aspectos - da ambigüidade que se observou no caso anterior.

Uma vez que se pode juntar o elemento lexical dessas palavras em uma mesma unidade lingüística (forma e conteúdo), o morfema que dá origem às formas ambíguas será único. A mudança de categoria ou de função sintática se deve à maneira como o morfema foi tratado pela língua, seu emprego e distribuição. A mudança de significado em parte se resolve pelos elementos e fatores que fizeram o morfema mudar de categoria e de função como no caso do substantivo *casa* e do verbo *casar*; e, em parte, pela maneira como a língua usa uma idéia - já definida na língua - para representar uma outra idéia com a qual a primeira tem alguma semelhança ou analogia, do ponto de vista com que a língua faz o recorte do mundo e como os representa através da linguagem. Assim a partir da palavra *casa*, moradia das pessoas, passou-se à idéia de *casa* como lugar em que ‘mora’ o botão, assim como da idéia de *casa*, moradia de pessoas, passou-se também à idéia de *casa de cachorro*, *casa de boneca*, *casa de carnes*, *casa de vinho*, etc. Os lingüistas chamam a esse fenômeno de ‘extensão do significado’. Quando ocorrer tal fenômeno, o morfema de origem estará presente em todos os demais das formas estendidas. As palavras resultantes representam itens lexicais diferentes somente na forma de superfície, sendo único na forma subjacente.

6. Alomorfia

Morfê é a representação fonológica de um morfema, sua forma morfológica. O morfema, constituído de forma e conteúdo, é uma unidade abstrata, uma representação obtida através da análise lingüística das línguas naturais. A maneira como essa unidade se atualiza na língua, ou seja, sua forma fonológica, chama-se morfê. Morfê é, pois, um conjunto de fonemas constitutivos de um morfema. Vimos anteriormente que, no vocábulo *indisponibilidade*, há dois morfemas que apresentam variação de forma, quando comparados com outros vocábulos em que ocorrem. Os elementos fonológicos que caracterizam essa variação são os morfês do morfema a que se refere. Neste caso, os morfês variantes são *po - pon* e *ível - ibil*. Os outros elementos formais da palavra também são morfês. O morfema inicial também apresenta variação de forma, dependendo do contexto em que aparece. Compare o *in-* de ‘indisponibilidade’ com o mesmo morfema com morfê diferente nas palavras *ilógico*, *inegável*. No primeiro caso,

temos a forma [iN], no segundo caso, temos a forma [i] e, no terceiro caso, a forma [in]. No conjunto, temos três morfes para um mesmo morfema¹⁴.

O fenômeno de variação dos morfes de um mesmo morfema chama-se alomorfia, por semelhança com 'alofonia', que é o fenômeno de variação de alofones de um mesmo fonema (cf. [s] e [z] em *rapaz* e *rapazes*).

A variação na forma de um morfema, ou seja, no seu morfe, representa um problema fonológico, ou seja, uma variação fonológica, se os elementos envolvidos podem ser estabelecidos através de regras e de um contexto fonológico, sem precisar fazer menção a noções morfológicas. Assim, o fato de o /S/ de plural realizar-se como [s] ou [z] depende do contexto fonológico em que ocorre: se estiver diante de pausa ou de consoante surda, ocorrerá como [s]; se estiver diante de vogal ou de consoante sonora, ocorrerá como [z]. Trata-se de uma regra fonológica e não morfológica, embora o /S/ (arquifonema de coda), neste caso, seja também uma desinência de plural (morfema de plural {S}).

A morfologia tem como questão fundamental a junção de forma e conteúdo. Em *casas* e *casas amarelas*, o -s final de *casas* tem duas pronúncias: [s] e [z]. As duas realizações representam um item lexical e são os morfes, ou seja, as realizações do morfema. Os dois morfes [s] e [z] são alomorfes do morfema {S} de plural. Portanto, o plural em Português apresenta *alomorfia*. Esse tipo de alomorfia, apresentado acima, é só fonológica, condicionada pelo contexto fonológico e, nem todos a reconhecem como alomorfia real. As alomorfias de raiz são casos de variação morfológica e não apenas fonológica. Nas palavras *tó-ra-k(i)s* e *to-rá-ci-co*, tirando o sufixo -ico, nota-se que a raiz varia na forma (não no conteúdo): portanto, apresenta alomorfia. Em *elétrico* e *eletricidade*, há alomorfia entre um morfe com [k] e outro com [s]: no primeiro caso, foi acrescentado o sufixo -ico e, no segundo caso, o sufixo -idade.

Silva & Koch (1987:30) dizem:

Tem-se, assim, casos de alomorfia fonologicamente condicionada: o morfema lexical gás apresenta uma variante gaso- em gasôênio e o morfema derivacional -eira apresenta uma variante -leira, em chaleira.

¹⁴ Nas transcrições, [] são usados para representar os segmentos fonéticos, // representam os fonemas ou arquifonemas, | | representam transcrição morfofonológica e { } representam o morfema. Neste trabalho, é comum os morfemas aparecerem apenas nas suas representações morfofonológicas.

Uma outra forma de interpretação desses casos considera que, em *gasômetro*, existe uma vogal de ligação -o- e, em *chaleira*, existe uma consoante epentética ou de ligação -l-. Desse modo, não haveria alomorfia nesses casos.

O caso do plural com morfes [s] e [z] pode ser simplesmente um arranjo fonológico e não um caso de alomorfia, segundo alguns gerativistas. Essa variação é diferente de variantes que trazem mudanças mais drásticas para os morfemas, do tipo *fazer* / *feito*, etc. Alguns autores chamam esse tipo de variação de substituição (*suppletion*). O caso *fazer/feito* (como 'go' e 'went', em Inglês) é de substituição total (*total suppletion*). Exemplos como *bom*, *melhor*, *ótimo* também poderiam ser considerados casos de substituição total.

Alguns casos de alomorfia (fonológica) representam formas que sofreram regras fonológicas diferentes em épocas diferentes da história da língua (cf. *leite* / *láceteo*) pelas semelhanças fonológicas entre essas formas, são chamadas de *partial suppletion*.

7. Tipos de morfemas

Apresenta-se, a seguir, uma lista de entradas com definições de termos e explicações de fenômenos que caracterizam diferentes tipos de morfemas, conforme são encontrados na literatura e usados em diferentes abordagens teóricas. Algumas noções apresentadas nesta seção serão desenvolvidas mais adiante, como a noção de flexão, derivação e composição, entre outras.

Morfema lexical é uma unidade morfológica que se agrega a outra pelo significado e não pela função sintática, como em *in* + *feliz*, em que há dois morfemas lexicais (ou morfes).

Morfema gramatical é uma unidade morfológica que se agrega a outra por causa de uma função sintática, como os morfemas de tempo, modo, pessoa, plural, etc. No verbo *achávamos*, *ach-* é um morfema lexical; *-a-* é a vogal temática (morfema gramatical) do verbo de 1ª conjugação; *-va-* é o morfema gramatical de tempo e modo (imperfeito do indicativo); *-mos* é o morfema gramatical de pessoa e número (1ª pessoa do plural).

Morfe é a realização fonética dos morfemas. Quem interpreta o morfema como tendo apenas uma informação abstrata, o morfê é a forma fonológica dos morfemas.

Morfemas relacionais são morfemas que ordenam os elementos da frase, fazem concatenação de morfemas lexicais, como as *preposições, conjunções e pronomes relativos*.

Morfemas 'bivalentes': os dêiticos são palavras que têm um morfema gramatical que traz consigo a noção básica da função dêitica (pronome demonstrativos, advérbios de lugar e de tempo, 1ª e 2ª pessoas verbais, etc.), e um morfema lexical 'vazio', cujo significado deve ser preenchido por informações extra-lingüísticas, ou seja, que vêm do contexto de situação do discurso. Podemos chamar esse tipo de morfema de 'bivalentes'. Assim, quando se diz: *eu nasci aqui*, o significado do pronome pessoal é completado, quando o interlocutor vê quem fala e o identifica como sendo fulano de tal. Se essa informação já tiver sido dada no discurso que precede esse enunciado, a função dêitica se confunde com a função anafórica. Com a palavra *aqui* além da função gramatical de advérbio de lugar, o significado desse morfema só se completa quando o interlocutor substitui o *aqui* pelo nome do lugar (significado lexical do morfema).

Morfemas flexionais são os morfemas que integram os paradigmas das flexões nominais e verbais. Os morfemas flexionais são também comumente chamados de 'desinências'.

Morfema flexional aditivo é aquele que acrescenta um morfema ao radical, como a formação do feminino *professor + a*.

Morfema flexional subtrativo é um morfema (*discutível*) que cai para mudar uma flexão, como acontece na passagem do masculino ao feminino da palavra *órfão = órfã*.

Morfema flexional alternativo é aquele que apresenta uma variação fonológica na qualidade vocálica de um segmento, como acontece com as formas *avô* e *avó* ou, ainda, nas alternâncias do tipo *formoso - formosa - formosos*). O caráter 'alternativo' dessa classificação é discutível.

Morfema zero / afixação zero: dada uma estrutura morfológica para um tipo de palavra, pode acontecer de ocorrer um ou

mais vazios na forma fonética da palavra. Esses vazios, que deveriam ser ocupados por desinências, comportam morfemas zero. Todo substantivo do Português tem uma estrutura que incorpora a definição de gênero e de número. Na palavra *casas*, o [S] final é o morfema de plural, e o [a] é o morfema de feminino. Na palavra *casa*, falta o morfema de plural e não há marca fonológica para o morfema de singular. Neste caso, o morfema de singular é um morfema zero. Nida (1949) defende a presença de um morfê zero, dizendo que a falta de uma marca morfológica é tão relevante quanto a sua presença, na formação das palavras. Alguns lingüistas chamam os morfemas zero de 'morfemas fantasma' (cf. em Inglês: *ghostly components*), como Hass (1960).

Um problema do morfema zero é que ele aparece 'demais', preenchendo muitos 'buracos' no sistema. Praticamente, em todo paradigma morfológico, encontra-se um morfema zero. Outro problema é o fato de ser um morfema 'mal definido', 'mal formado' e de uso 'irrestrito'. A partir de um exemplo, é difícil, senão impossível, dizer qual a natureza e função de um morfema zero.

Morfema latente é o mesmo que morfema zero.

Morfema fantasma: em alguns casos, um morfema que ocorre na forma subjacente desaparece durante a derivação, mas deixa um traço que serve de gatilho para a aplicação de alguma regra de reajuste morfofonológico. Esse morfema que desaparece na derivação é chamado de morfema fantasma.

Morfema livre e morfema preso: um morfema livre é aquele que pode ser uma palavra por si. Por exemplo, a preposição *com* é um morfema livre. Um morfema livre corresponde a um radical (ou raiz). Um morfema preso representa uma unidade morfológica que sempre vem grudada a outra, como os afixos. Alguns morfemas presos têm um significado de difícil definição, sendo mais clara sua função sintática, como acontece com as desinências.

Forma livre / presa é o mesmo que morfema livre / preso.

Afixo: todo morfema que não é raiz (ou radical) é um afixo. O termo afixo é abrangente, referindo-se preferencialmente aos prefixos e aos sufixos, mas podendo incluir também desinências, em geral. Os morfemas gramaticais são sempre afixos. Nas palavras compostas, pode-se encontrar dois ou mais morfemas lexicais que não são afixos,

como em *cirurgião-dentista*. Na palavra *reciclável*, antecedendo o morfema lexical *-cicl-*, ocorre o afixo *re-* (prefixo) e, no final, o afixo *-avel* (sufixo).

Prefixo é o morfema que precede a raiz ou radical. Uma palavra pode ter mais de um prefixo, como em *mididesvalorização* em que se encontram os prefixos: *midi* e *des*. Alguns autores acham que os prefixos são morfemas independentes, não afixos.

Sufixo é o morfema que ocorre depois da raiz ou do radical. Um tipo especial de sufixo é o morfema que funciona como uma flexão. Neste caso, costuma ser chamado de desinência e não de sufixo.

Terminação é um termo que se refere a tudo o que aparece depois do morfema lexical, englobando os sufixos, as vogais temáticas e as desinências, como grupo ou com referência a um deles.

Desinência é todo morfema que pertence a um paradigma de elementos gramaticais, como os morfemas de gênero, número, tempo, modo, pessoa, etc. As desinências sempre são morfemas gramaticais. O conjunto de desinências que formam um paradigma é chamado de flexão. Em Latim, os nomes tinham flexão exemplificada pelas declinações: *rosa, rosae, rosam, rosarum, rosas*, etc. Em Português, a flexão dos nomes ficou restrita ao paradigma do gênero e do número dos nomes. Para alguns autores, os paradigmas de grau também formariam uma flexão (estendível também aos advérbios, em certos casos). Para outros, as terminações que marcam o grau são sufixos comuns. Sandmam (1991: 33) diz "... o morfema de grau não é mais visto como uma flexão". Mattoso Câmara (1972: 71) já tinha se posicionado do mesmo modo.

Há, ainda, quem distinga casos de derivação (como o diminutivo com *-inho*) de casos de composição, ou mesmo, de ocorrência de duas palavras distintas, com o diminutivo *-zinho*. Neste último caso, uma palavra como *casinha* teria uma realidade morfológica diferente de uma palavra como *florzinha*. No caso dos verbos, tanto em Latim como em Português, ocorrem os paradigmas verbais que dão origem às conjugações verbais.

As desinências ocorrem sempre no final das palavras e depois dos sufixos, quando estes estiverem presentes, como em *cas-inh-a-s*; *rítm-ic-o-s*; *suav-iz-á-sse-mos*, em que as desinências aparecem depois dos sufixos *-inh-*, *-ic-*, *-iz-*.

Conjugação é o conjunto de morfemas gramaticais que formam as flexões verbais de uma língua.

Declinação é o conjunto de morfemas gramaticais que formam as flexões nominais.

Paradigma é o conjunto de flexões que exerce uma mesma função. Em um sentido mais amplo, usado pela Linguística, entende-se por paradigma qualquer conjunto de elementos que exerce uma mesma função.

Infixo é o morfema que aparece inserido dentro de outro, como os exemplos de mesóclise em Português: *dar-te-ei* (cf. *te darei*). O plural da palavra *qualquer* é *quaisquer*. Na forma plural, a desinência de plural *-s-* é um infixo. Nos compostos do tipo *cirurgiões-dentista*, a primeira palavra flexiona em gênero e número *-ões*. Como se trata de um item lexical único, os morfemas de gênero e de número deveriam ser considerados infixos. Contudo, tradicionalmente, não o são, sendo apenas desinências do primeiro lexema ou da primeira palavra do composto. Historicamente, como o futuro dos verbos em Português veio de uma forma composta 'verbo principal' + 'verbo *haver*', os casos de mesóclise eram (ou ainda são?) do mesmo tipo dos substantivos compostos, referidos acima. Comparem-se, ainda, as formas *dar-te-ei* com *hei de te dar*. No segundo caso, como acontece com os verbos compostos, a inserção do clítico não é considerada um caso de infixação.

Vogal temática é um tipo especial de desinência (ou 'sufixo'), formado por uma vogal, que define um paradigma flexional, funcionando como uma espécie de rótulo (Lee, 2000). Os nomes são identificados pelo fato de apresentarem uma terminação vocálica que remete ao paradigma de gênero. Por essa razão, alguns autores, como Mattoso Câmara (1970), consideram essas vogais como sendo vogais temáticas nominais, além de morfemas de gênero (Lee, 2000). Nesse sentido, costuma-se dizer que *-o* refere-se ao masculino; *-a*, ao feminino e *-e* refere-se a um gênero definido como 'neutro' (nem masculino, nem feminino). Na verdade, a questão não é tão simples assim e a explicação acima parece referir-se mais à escrita do que ao sistema da língua. Uma vogal temática é um morfema classificatório, indicativo de uma categoria.

No caso dos verbos, as vogais temáticas classificam-nos em conjugações: a 1ª conjugação tem a vogal temática *-a-*: *cant-á-ssemos*. A 2ª conjugação tem a vogal temática em *-e-*: *vend-e-mos*; a 3ª conjugação, vogal temática em *-i-*: *part-i-amos*.

Composição é o processo de formação de compostos e pode ser de dois tipos: aglutinação e justaposição.

Justaposição é o processo de composição, em que os elementos conservam sua forma original completa: *passatempo*.

Aglutinação é o processo de composição em que ocorre uma fusão das palavras, resultando em um todo único, com só um acento principal. Em geral, ocorre alguma mudança fonológica na forma final da palavra. Palavras como *boquiaberto* e *pontiagudo*, citado às vezes, não são bons exemplos: primeiro, porque podem ter mais de um acento principal; em segundo lugar, parece ser o caso de uma justaposição com alteração fonológica. No exemplo *planalto* ocorre uma verdadeira aglutinação. Na aglutinação, o padrão prosódico acentual é o mais importante.

Cruzamento vocabular, (em Inglês: *blends*; em Francês *portmanteaux*) é o processo de composição, formado por duas palavras que perdem parte de seus elementos, fundindo-se em uma unidade fonológica e prosódica, portadora de um único acento principal. Em Inglês, das palavras *channel* + *tunnel* originou-se a palavra amalgamada *chunnel*. Em outras palavras, amalgama é a junção de partes de duas palavras para formar uma nova palavra. Trata-se de um processo muito popular, porém, que costuma formar palavras de vida curta na língua, sendo, em geral, um modismo de ocasião. Raras são as palavras que acabam se incorporando definitivamente ao léxico da língua. Exemplos: *Goianobyl* (Goiânia + Chernobyl).

Compostos são morfemas formados por duas ou mais raízes, tendo dois radicais, carregando um significado próprio, diferente da soma dos significados dos morfemas que o compõem. Exemplo: *guarda-chuva* (objeto). Os compostos mantêm a categoria morfológica do último morfema. Vários tipos de categorias sintáticas podem formar sintagmas tidos como uma unidade morfológica única.

Composto copulativo é um substantivo ou adjetivo composto, formado por palavras que especificam o significado da anterior,

dentro da mesma área semântica, como em *cantor-compositor*, *biofísicopsicossocial*, *hortifrutigranjeiro*; *meia-calça*, *nipo-brasileiro*; *italo-brasileiro*; *afro-brasileira*. São chamados também de compostos coordenados.

Compostos clássicos são morfemas formados por raízes gregas ou latinas que funcionam como se fossem raízes independentes (embora possam parecer afixos) e formam um certo tipo de compostos. Exemplos: *hiper-mercado*; *hiper-inflação*; *hiper-legal*.

Hibridismo são compostos formados por raízes de línguas diferentes, como: *automóvel* (Grego com Latim), *sociologia* (Latim com Grego), *bananal* (Taíno - das Índias Ocidentais - com Latim).

Etimologia é a relação que uma palavra tem com outra mais antiga da língua, da qual se derivou. A palavra de origem chama-se também *étimo*. Entre a palavra de origem e a atual, os estudos diacrônicos (e filológicos) mostram que houve a aplicação de uma série de processos gramaticais, sobretudo fonológicos e morfológicos, resultando na forma e função atuais da palavra.

Arcaísmo é uma forma lexical, uma construção ou um significado, cujo uso não ocorre mais na língua, mas que era comum em outro momento da história da língua, como *u* significando *onde*.

Forma residual ou fóssil é uma forma fonológica de morfemas que conserva características fonéticas, morfológicas, sintáticas ou semânticas próprias de um estágio antigo da língua, tido como ultrapassado (cf. *entonce*, hoje *então*).

Empréstimo lexical: palavra de origem estrangeira usada com adaptação fonológica, porém, nem sempre com adaptação ortográfica. Exemplos: *blush*, *stress*, *input*, *ateliê*, *butique*, *freezer*, *show*, *contêineres*, *shopping*.

Palavras inventadas (cf. em Inglês *word manufacture*), como o rótulo indica, são palavras inventadas arbitrariamente por pessoas e que passaram ao domínio comum. Muitos nomes de produtos acabam se tornando palavras mais abrangentes do que indicava seu significado original. Um *xerox* pode ser uma fotocópia, sem ter sido feita por uma máquina da marca Xerox.

Neologismo é uma palavra nova, inventada para designar uma realidade nova. Os neologismos são formados a partir do estoque de morfemas da língua e não apenas a partir do estoque de fonemas.

Substratos são marcas de uma língua que foram assimiladas por outra. Os substratos referem-se somente a elementos sonoros e morfológicos.

Amalgama é o mesmo que cruzamento vocabular.

Derivados são morfemas de apenas uma raiz, que mudam de categoria morfológica ao receber ou perder sufixos ou desinências. Da palavra simples *água* tem-se o verbo derivado *aguar*. Do verbo simples *cortar* tem-se o substantivo derivado *corte*. O adjetivo simples *belo* forma o substantivo derivado *beleza*.

Incorporação: em algumas línguas, como o Chukchee (Spencer, 1991: 15), o verbo e o objeto (direto ou indireto) formam um todo morfológico único. Em Chukchee, a palavra composta [tə-qora-pelarkən] é formada por verbo mais objeto, mantendo suas funções sintáticas (cf. *tə* = eu; *pelarkən* = deixar, deixando; *qoraŋə* = a rena). Em outras línguas, a maneira como a harmonia vocálica se estende pelos limites de uma palavra pode mostrar que certos sintagmas foram ‘incorporados’ morfofonologicamente e valem por palavras simples.

Circunfixos são as duas partes descontínuas de um único morfema, isto é, um morfema é interrompido por uma ou mais palavras intercaladas, como acontece com o morfema de negação *ne ... pas* do Francês (je *n'ai pas* achetté le livre). Essas formas podem ser consideradas também como prefixos e sufixos agregados à palavra da qual dependem, como no caso da negação em Francês. Em Português, não há caso de ocorrência de circunfixos.

Reduplicação: em algumas línguas, uma parte inicial de alguns morfemas se repete. A morfologia considera esses casos como diferentes de prefixos, dando-lhes o nome de reduplicação. Em Grego, no aoristo, os verbos apresentam esse fenômeno (*lélēka*). As reduplicações marcam função sintática própria. Vistas apenas do ponto de vista fonológico, palavras como *Zezé*, *Lulu*, *papai*, *titio*, etc., podem ser interpretadas como um caso especial de reduplicação.

Abreviatura inicial (cf. Inglês *clipping*) é um processo de redução de uma palavra a uma ou duas sílabas iniciais, nem sempre coincidindo com a segmentação em morfemas. A Língua Inglesa usa muito esse processo, sobretudo no registro oral. Na escrita, há muita restrição com relação ao uso dessas formas, a não ser naqueles casos que já se tornaram consagrados pelo uso. Em Inglês, da palavra *laboratory* aparece a forma *lab*, de *telephone*, aparece a forma *phone*. A primeira é típica da linguagem oral e não da escrita, mas a segunda aparece tanto na fala, quanto na escrita. Exemplos do Português: *extra* (-ordinário), *cine* (-ma), *foto* (-grafia), *inox* (-idável).

Braquissemia é outro nome para ‘abreviatura’.

Acronímo é a formação de palavras a partir da junção das letras iniciais da escrita, como nos nomes ONU (Organização das Nações Unidas), OTAN (Organização do Tratado do Atlântico Norte). Um acronímo pode ser produzido também com finalidade de esconder ou camuflar uma expressão, como recurso estilístico. Por exemplo, em Inglês, pode-se usar a forma *misty* como acronímo de *more ideologically sound than you*. O famoso *Febeap* de Stanislau Ponte Preta é um acronímo que quer dizer *Festival de besteiras que assola o país*.

Anagrama consiste na inversão total das letras de um nome, de tal modo que resulte um outro: *Ronoel* = *Leonor*; *Belisa* = *Isabel*; *Iracema* = *América*; *Acemira* = *Iracema*.

Sigla é uma palavra formada da abreviatura de outras. Há vários processos de formação de siglas: algumas se servem apenas das letras iniciais, como *IBGE* (Instituto Brasileiro de Geografia e Estatística); outras pegam as palavras principais, como *PT* (Partido dos Trabalhadores), *RS* (Rio Grande do Sul); outras unem pedaços das palavras originais, como em *UNICAMP* (Universidade Estadual de Campinas), *PUCAMP* (Pontifícia Universidade Católica de Campinas). As siglas admitem a derivação, como em *petista*, *unicampense*, etc.

Hipocorístico é o processo usado na linguagem familiar para traduzir carinho ou qualquer palavra criada por afetividade, incluindo diminutivos e palavras da linguagem infantil, como as alterações do prenome ou do sobrenome (apelidos): *José* - *Zé*, *Zezinho*, *Zeca*, *Zeza*,

Zuza; Anacleto - Cleto; Benedito - Benê; Eptácio - Pita; Albertina - Berta; Francisca - Quica; Filomena - Filó; Carlos Eduardo - Cadu.

Cisão morfológica ocorre quando um morfema adquire um significado próprio, diferente, do que apresenta em outros contextos. Essa cisão pode levar à formação de formas divergentes (cf. em Inglês *person* e *parson* 'pessoa' e 'pároco').

Formas divergentes aparecem quando ocorre uma diferenciação fonológica de um morfema, gerando novo morfema com a adição de novo significado.

Cognatos são morfemas que compartilham o uso de uma mesma raiz. A partir de uma base etimológica comum, em línguas de uma mesma família, através do processo de evolução divergente, podem-se formar morfemas chamados cognatos, porque apresentam um grau estreito de parentesco lingüístico.

Fusão morfológica é o resultado da anulação de uma distinção morfológica de uma forma fonológica determinada, como aconteceu na perda dos casos ao se passar do Latim ao Português.

Contaminação: por semelhança de pronúncia, de significado ou de função, uma forma pode interferir com outra historicamente na evolução da língua, mas a forma contaminada não perde sua identidade.

Cruzamento: duas formas morfológicas semelhantes (fonológica, semântica ou sintaticamente) que, na evolução da língua, são interpretadas como formas em competição.

Pseudo morfema: em *descascar*, *desobedecer*, a segmentação é fácil, mas o significado do *des* não é o mesmo. Em *distúrbio*, o *dis* é um pseudo morfema, uma vez que não existe *túrbio*, na língua... e *distúrbio* é um 'morfema simples e único'.

Formas alternantes: elementos de uma alomorfa. Em geral, são formas condicionadas gramaticalmente ou morfológicamente, como as formas do participio passado em Inglês.

Nominalização: processo morfológico que agrega morfemas (sufixos) para mudar a categoria de uma palavra (verbo ou adjetivo),

tornando-a um substantivo. Em Português, os principais sufixos nominalizadores são *-ança*, *-ção*, *-mento*, *zero*. Exemplos: *lembrança*, *nomeação*, *sofrimento*, *a pesca*.

Onomatopéias são palavras que procuram imitar sons da natureza, como *zumzum*, *tique-taque*, etc.

Oneônimos: são processos morfológicos formadores de nomes comerciais, como em *mentex*, *eucalol*, etc.¹⁵

8. Estrutura dos nomes e dos verbos

8.1. Classes de palavras

Classe de palavra é o termo tradicional usado para se referir às categorias morfológicas da gramática, do tipo substantivo, adjetivo, verbo, etc. Categoria lexical é o mesmo que classe de palavra.

Classe morfológica é o conjunto de palavras que recebem um tipo de flexão e que exercem uma mesma função, por exemplo, os adjetivos masculinos recebem a flexão com a desinência *-o* e formam uma classe morfológica. Todas as palavras marcadas morfológicamente como masculinas formam a classe morfológica das palavras do gênero masculino. Uma classe é um conjunto mais ou menos abrangente, dependendo do ponto de vista como é definida.

Classe lexical é o conjunto de palavras, às quais se aplica uma determinada regra, por exemplo, de gênero, de tipos de verbos (transitivos, intransitivos), etc. O termo 'classe lexical' equívale à 'classe morfológica'.

8.2. Clíticos

Clítico é uma palavra que, embora tenha uma identidade morfológica própria, aparece sempre grudada sintaticamente em outra, chamada de 'hospedeira'. Os pronomes oblíquos do Português são exemplos de clíticos, pois aparecem sempre ligados a um verbo; por exemplo: 'eu *te* vi'; 'Maria *lhe* disse'; 'achei-os na gaveta', etc. Na expressão *ele mo disse* há dois clíticos: *me + o* (cf. **ele me o disse*). Os falantes podem dizer *tu!*, mas não podem dizer **te!*. Embora não

¹⁵ O termo *oneônimos* (formado do Grego *onéo* = *comprar* e *ônimo* = *nome*) parece ter sido lançado por Mansur Guérios.

possam dizer *ti!*, podem dizer *para ti!* (cf. **para te!*). Assim, *te* é um clítico, porém, *ti* não o é. A caracterização dos clíticos entram em jogo também aspectos prosódicos: os clíticos são, por natureza, átonos. Em outros termos, pode-se dizer que os clíticos combinam-se fonologicamente com palavras das quais dependem prosodicamente, mas não formam uma unidade morfológica única. Desse modo, os clíticos distinguem-se também das palavras compostas. Os clíticos grudam-se a verbos ou a substantivos. Os artigos são clíticos dos substantivos, mas outros determinantes, como os pronomes demonstrativos, não são clíticos. Em enunciados como *a casa de Júlia é antiga*, a preposição *de* é um clítico, mas na expressão *a casa é antiga para mim, não para você*, a preposição *para* não é clítico, porque pode levar acento principal. Ao contrário do que diz a gramática tradicional, as preposições com mais de uma sílaba não são átonas por natureza.

Proclítico é um clítico que precede a palavra hospedeira na cadeia da fala.

Enclítico é um clítico que ocorre logo após a palavra hospedeira na corrente da fala.

8.3. Consoante / vogal de ligação

Em alguns casos a junção de morfemas requer a presença de uma consoante ou de uma vogal para que o contexto fonológico não fique mal formado. Esses elementos fonológicos, não portadores de um significado próprio (portanto, não são morfemas) são chamados de consoante ou de vogal de ligação. Ocorrem sempre na junção intervocabular. A epêntese refere-se à inserção de um ou mais fonemas (consoante ou vogal) dentro de uma palavra ou fazendo a junção entre palavras ou morfemas. Para quem considera que o diminutivo do Português é *-inho*, a forma com *-zinho* apresenta uma consoante de ligação *-z-* (caso de epêntese). Para quem considera que o morfema de plural é apenas *-s*, os plurais em *-es* têm uma vogal *-e-* de ligação, como em *professor-e-s*; *mal-e-s*; *rapaz-e-s*. Convém observar que esses segmentos ocorrem sempre em junção morfológica, sendo diferentes dos que ocorrem em outros contextos, por exemplo, para completar a formação de uma sílaba, como o *-i-* de [xapa-i-s] (*rapaz*), [vose-i-s] (*vocês*).

Consoante / vogal de preenchimento é um caso semelhante ao da consoante / vogal de ligação, porém, que ocorre dentro de palavras. Uma palavra como *sonho* pode ter um ditongo nasalizado em vez de uma vogal simples nasalizada. Na palavra *muito*, após o ditongo

nasalizado, pode ocorrer uma consoante nasal palatal. A inserção desses elementos não muda a estrutura morfológica, mas apenas a fonológica e a aparência fonética das palavras.

8.4. Raiz, Radical, Tema

Raiz é um morfema lexical; é o morfema que resta quando todos os afixos são removidos. Nem toda raiz tem um morfema que pode ocorrer isoladamente. Em Inglês, o morfema raiz *dent* forma palavras como *dentist*, *dental*, *dentures*, mas não há uma palavra *dent* que se possa usar isoladamente, igual a *tooth*. Em Português, há uma raiz *dent* da qual se derivam as palavras *dental*, *dentadura*, *dentista*. A vogal *-e* final é uma vogal temática de gênero (neutro ou masculino - dependendo da abordagem teórica empregada).

Existe um uso diferente do termo 'raiz' (cf. em inglês *root*). Na tradição lingüística portuguesa, sobretudo da gramática normativa e da filologia, o termo *raiz* é usado para designar apenas os morfemas das línguas que deram origem aos modernos morfemas da língua. Em outras palavras, na Língua Portuguesa, a maioria das palavras é formada por morfemas, cuja origem é uma 'raiz' latina ou grega. Uma palavra como *futebol* é formado por duas raízes inglesas: *fit* + *bol* (cf. em inglês, *foot* + *ball*). A palavra *abajur* tem duas raízes francesas (em Francês: *abat-jour*). Na palavra *super-mercado*, o morfema *super* é uma raiz latina; porém, na palavra *hiper-mercado*, o morfema *hiper* é uma raiz grega. Na literatura lingüística em Língua Inglesa, o termo *root* é empregado para designar o morfema lexical básico das palavras, o que sobra depois que todos os afixos foram removidos. Para Hockett (1954), *a simple form is a root, a derived form is an underlying form to which a process has been applied* (uma forma simples é uma raiz, uma forma derivada é uma forma subjacente à qual foi aplicado um processo – Malmkjaer, 1991: 323). Matthews (1974: 39-40) não gosta do termo *root*. Malkiel (1978) sugeriu o uso do termo *primitives* ('primitivos') em lugar de *root*.

Para Basílio (1989: 91), *raiz é o elemento mínimo constitutivo da palavra que pode, por si só, constituir uma base ou mesmo uma palavra; núcleo mínimo de um radical ou base*.

Raiz livre é uma raiz que pode ser usada como palavra independente, de forma isolada, como *extra*, *super*, *hiper*, *abajur*. O contrário é uma raiz presa. Os verbos têm sempre suas raízes lexicais presas, assim como os nomes que têm desinências de gênero e de número. Os morfemas dos sufixos e muitos dos prefixos são formados por raízes presas.

Monteiro (1991: 12) refere-se ao termo semantema, dizendo que é a parte da palavra em que repousa a significação lexical básica. Constitui o que se denomina comumente de raiz.

Radical é uma unidade formada pela raiz e pelos afixos, deixando de lado as desinências. A partir de uma forma de base ou raiz lexical, pode-se ter radicais formados pelo acréscimo sucessivo de afixos, como em *mar*, *mar(es)*, *marinh(a)*, *marinheir(o)*, *marítim(o)*, *submarín(o)*. Nesses exemplos, nota-se claramente a diferença que existe, na tradição portuguesa, entre raiz e radical. A raiz é apenas *mar*, mas dessa raiz podem-se formar muitos radicais, bastando para isso remover as flexões e as vogais temáticas. O radical corresponde à raiz somente na forma mais primitiva da palavra. O termo inglês que corresponde à noção de radical é *stem*, ou seja, é a formação morfológica que agrega todos os elementos, exceto as desinências de flexão.

M. McCarthy (1991: 318) dá o seguinte exemplo de análise morfológica para *repainted*:

root	<i>paint</i>
affixes	<i>(re-) paint (-ed)</i>
stem	<i>repaint (-ed)</i>
morphs	<i>re-paint-ed</i>
morphemes	AGAIN + PAINT + PAST

O núcleo de uma construção morfológica recebe o nome de *base* (em Inglês: *head*). É a forma sobre a qual atuam os processos de formação de palavras (Basílio, 1989: 90).

Tema é o radical acrescido de vogal temática (cf. flexão).

9. A flexão

Flexão é um conjunto de elementos da mesma natureza, ou seja, um paradigma, que agrega elementos de uma mesma categoria sintática. Os morfemas *-o* e *-a*, que representam o masculino e o feminino em Português, agem morfológicamente através de um processo chamado de flexão de gênero.

Algumas línguas têm um processo de flexão muito rico e variado, como o Latim e o Grego, quer para os nomes quer para os verbos. O Português tem um sistema verbal rico em flexões, mas com relação aos nomes, tem apenas as categorias de gênero e número.

Alguns autores incluem a categoria de grau. O paradigma das flexões chama-se também **declinação**.

Sandman faz a seguinte observação:

“No verbo, a conjugação (1ª em -a, 2ª em -e e 3ª em -i) e o pertencer à classe regular ou irregular são traços lexicais, porém sem conseqüências semânticas. As flexões de tempo, modo e aspecto, traços gramaticais, são cumulativas e têm valor semântico, fazendo parte da estrutura profunda. As de número e pessoa, traços gramaticais também e cumulativas, são fruto da concordância com o sintagma nominal-sujeito, não têm valor semântico e são introduzidas na estrutura de superfície. Observe-se, a propósito, que na variante inculta do português as flexões de número e pessoa do verbo foram totalmente ou em grande parte suprimidas: ‘eu trabaio, tu, você ou vancê trabaia, ele trabaia, nós trabaia, voceis ou vanceis trabaia, eles trabaia...’ (Sandman, 1991: 41-42).

As desinências de flexão se acrescentam no último estágio da formação das palavras, depois que os demais afixos já foram agregados (depois que todos os processos derivacionais e de composição já foram realizados)

9.1. A flexão de gênero

Todos os autores admitem que, em Português, uma das funções do artigo é indicar o gênero a que pertencem os nomes (substantivos e adjetivos). Em alguns casos, a presença do artigo revela o gênero do nome de maneira mais clara, como em *o dia* (esperar-se-ia que uma palavra terminada em *-a* fosse feminina)¹⁶; *o dentista* e *a dentista*, caso em que a não presença do artigo revela uma palavra ambígua com relação ao gênero.

Algumas palavras alternam formas masculinas com femininas, com a simples mudança da vogal fina, a qual tem que ser ‘a’ (feminino) ou ‘o’ (masculino), como em: *peru - perua*, *autor - autora*, *pombo - pomba*. Recentemente, tem notado com maior freqüência um tipo de alternância com ‘e’ (neutro / masculino), como em *parente - parenta*, *governante - governanta*.

A maioria das palavras são ou masculinas ou femininas de acordo com sua função lingüística. Convém lembrar que a idéia de

¹⁶ É interessante a palavra *o cara* para se referir a pessoas do sexo masculino, não ocorrendo uma forma **a cara* para as mulheres. Todavia, sem o artigo, pode ocorrer *cara* para homem ou para mulher.

gênero não é igual à idéia de sexo. Em uma frase como *Pedro é uma pessoa amiga*, a palavra *amiga* atribuída ao Pedro está no feminino porque tem que concordar com a palavra *pessoa*. Quando se trata de animais, a língua procurar distingui-los, mantendo uma relação mais próxima com a realidade biológica, embora isso nem sempre aconteça. A palavra *criança* é feminina, mas há as palavras *menino* e *menina*. A palavra *homem* é masculina na Língua Portuguesa, e *mulher* é uma palavra feminina. Do ponto de vista lingüístico, *mulher* não é uma forma feminina de *homem*, mas uma palavra independente.

Os nomes em Português podem terminar também pela vogal ‘e’.¹⁷ Essas palavras são consideradas por alguns lingüistas como formando uma espécie de gênero ‘neutro’ - nem masculino, nem feminino. Como, porém, podem ser precedidas pelo artigo, e este é ou masculino ou feminino, não faz sentido dizer que existe um gênero neutro em Português.

Alguns pesquisadores chamam de vogal temática ou de marcador de classe de palavra, a vogal que acima foi interpretada como sendo flexão de gênero. O mais correto seria atribuir a essas vogais a dupla função de marcar o gênero e a classe de palavra. Nos nomes, portanto, a vogal temática se confunde com a vogal de gênero. Silva e Koch (1987: 41) concluem que *conseqüentemente, a flexão de gênero nos nomes é um traço acessório, redundante*. Segundo Sandmam (1991: 39), o adjetivo tem flexão de gênero, porém, o substantivo tem sufixo derivacional: *bonito, bonita*. Citando Mattoso Câmara (1970: 78), esse autor lembra que, quando há substantivos que alternam o masculino com o feminino, como em *o barranco - a barranca* o feminino denota uma especialização do masculino. Vejam-se, ainda, os seguintes exemplos: *jarra - jarro; barco - barca*.

As palavras que terminam foneticamente por consoante, oriunda dos arquifonemas /L, R, S, N/, apresentam um morfema zero de gênero. Exemplos: *o algoz, a paz, o caráter, o mártir, a mártir, o homem, o mal, a cal*.

Como foi visto anteriormente, algumas formas alternam o gênero pela presença ou ausência do elemento vocálico final, como em *órfão - órfã, irmão - irmã, réu - ré, mau - má*. Outras alternam a qualidade vocálica, como em *avô - avó, formoso - formosa - formosos*.

¹⁷ O uso da escrita ‘a’, ‘o’ e ‘e’ corresponde às terminações vocálicas /a/, /u/ e /i/. Os ditongos acabados em [w] também tendem a ser masculinos: *sol, mal, fuzil* (cf., no entanto, *a cal*). Os ditongos terminados por [j] também tendem a ser masculino: *pai, herói*. Essa é uma das razões pelas quais o plural de algumas palavras terminadas por [w] se faz trocando o [w] por [i], fato que mantém o mesmo gênero: *farol - faróis, carnaval - carnavais*.

Algumas terminações também têm a função de mudar o gênero dos nomes, como se pode observar nos seguintes exemplos: *-isa, -essa, -esa: diaconisa (diácono); abadessa (abade); baronesa (barão)*, ou ainda em: *perdigão - perdiz; imperador - imperatriz*. Essas terminações são tradicionalmente consideradas sufixos, mas, de fato, integram também o paradigma de flexão de gênero da língua. Por outro lado, é preciso não confundir os casos acima com variações fonológicas do tipo *valentão - valentona; europeu - européia; judeu - judia*. A terminação *-inha* de *galinha*, certamente, não é um formativo de feminino de *galo*, nem de diminutivo. A variação de gênero nessas duas palavras deve-se à terminação *-o* e *-a* (além do tipo de artigo que elas podem receber).

9.2. A flexão de número

A categoria de número, na gramática da Língua Portuguesa, marca palavras com referência a mais de um referente da mesma natureza. A palavra *banana* refere-se a uma só banana e a palavra *bananas* refere-se a mais de uma banana. Nem toda idéia de plural no mundo tem uma marca de plural na linguagem, como acontece com os coletivos, que encerram mais de um elemento formando um conjunto e cujo plural refere-se apenas ao conjunto e não aos elementos desse conjunto, como em *cardume (muitos peixe reunidos)* e *cardumes* (vários conjuntos de ‘muitos peixes reunidos’).

Do ponto de vista de como a linguagem representa o mundo, nota-se também que há palavras na forma plural que remetem a objetos singulares, como *óculos, exéquias, núpcias*, etc. Atualmente, algumas palavras apresentam uma forma também no singular, que não existia antes, como *calça* (de *calças*), *algema* (de *algemas* - cf. *prende a algema da direita no corrimão*).

A flexão de número, quer nos nomes quer nos verbos, é marcada pelo morfema [S], que ocorre após as desinências de gênero, sendo sempre o último elemento à direita das palavras. Em alguns casos, ocorre uma vogal epentética [i], como em *paz - pazes, país - países, professor - professores, mal - males*. A razão disso está na maneira como a língua trata a formação de codas, não admitindo certos grupos de consoantes.

Já foi observado anteriormente que palavras que terminam por [w], oriundo de um /L/, alternam o morfema de gênero [w] ~ [j], antes de receber o morfema de plural: *lençol - lençóis, fuzil - fuziis - fuzis, estrutural - estruturais*, etc.

À semelhança com o que aconteceu com algumas palavras, no caso da flexão de gênero, pode ocorrer variação na qualidade vocálica de algumas palavras, quando passam para a forma de plural, como em *corpo - corpos, morto - mortos*.

As palavras que alternam a terminação *-ão* com *-ões* tem no *-s* um morfema de número, e na vogal *-o-*, a flexão de gênero. A vogal *-e-* é, ainda, uma vogal epentética. A vogal *-ã-* pertence a um morfema de grau aumentativo, que se torna opaco no plural.

A marcação de plural nos nomes compostos segue, de certo modo, uma lógica: se as duas palavras podem traduzir a idéia 'natural' e gramatical de plural, ambas ficam com as marcas de plural, caso contrário, ficará no plural somente a que puder ou, então, o último elemento. Veja os seguintes exemplos: *altos-relevos, surdos-mudos, cirurgiões-dentistas, tenentes-coronéis, navios-escola, escolas-modelo cadeiras-de-rodas, gato-de-botas, luso-brasileiros, saca-rolhas, vira-latas, tico-ticos, esconde-escondes, treme-tremes, (os) bota-fora, (tecidos) azul-marinho, pontapés, planaltos*.

9.3. A flexão de grau

Nem todos os autores consideram a categoria de grau como pertencendo ao processo de flexão. Uma das características da flexão é seu caráter abrangente na língua e, nesse sentido, o que ocorre com os graus não é diferente, uma vez que praticamente todos os nomes podem ter a forma do aumentativo ou do diminutivo. Vendo a língua desse modo, nota-se claramente que o Português difere bastante do Latim, que é uma língua que tem a flexão de grau bem definida com desinências específicas. Em Português, *maior, menor, melhor e pior*¹⁸ não têm mais a desinência de grau, como acontecia em Latim. São palavras simples. A idéia semelhante à *mais grande, mais pequeno, mais bom e mais mau* fazem dessas palavras sinônimos perfeitos.¹⁹

As palavras sem as desinências de grau são consideradas palavras de grau normal. Algumas desinências (ou sufixos) formam os aumentativos: *livro - livrão, mulher - mulherona*, etc.; outras formam o diminutivo: *livrinho, livresco, caixote*.

¹⁸ Veja, ainda, as palavras *anterior, posterior, inferior, superior*. Fora essas palavras, que vieram diretamente do Latim, não se pode mais formar palavras novas com uma terminação de grau *-ior*, à moda do Latim.

¹⁹ Existe um preconceito das gramáticas tradicionais com relação ao uso dessas formas que tem feito com que elas não sejam usadas, por exemplo, na linguagem escrita.

Alguns autores acham que o morfema *-zinho* e outros iniciados pela consoante /z/, tradicionalmente considerados morfemas de grau, são itens lexicais autônomos, ao passo que *-inho* e outros semelhantes são desinências (ou sufixos) de grau (veja Lee, 1992).

Nota-se que as desinências de gênero e de número, que ocupam tipicamente o final das palavras, podem ocorrer 'dentro' de palavras com formação de grau: *cãezinhos, cãezitos, fogõezinhos*, etc.

A categoria de grau comporta ainda as formações do comparativo de superioridade e de inferioridade, contrapondo-se a um grau normal ou sem comparação. As formas *maior, melhor, menor e pior* eram marcadas em Latim com o grau comparativo de superioridade, equivalentes a 'mais + adjetivo'. Estas últimas eram consideradas 'relativas', pois exigiam um segundo termo na comparação. As formas superlativas eram absolutas e não exigiam um termo de comparação. Todas essas formas passaram para o Português com a mesma função e com a desinência *-issim(o,a)*: *belíssimo, belíssima*. Dada a situação atual da língua, seria possível dizer algo como *melhorzíssimo*, o que seria completamente estranho em Latim.

A função de grau, em Português, atingiu também alguns advérbios, não ficando restrita aos adjetivos, como nos exemplos: *devagarinho, devagaríssimo*.

10. A derivação

Derivação é um processo morfológico que soma morfemas a um radical, fazendo com que a palavra resultante mude de categoria sintática, além de receber algum significado, carregado pelo próprio morfema. Alguns sufixos fazem com que substantivos tornem-se adjetivos (aglutinação de morfemas).

O processo de derivação aceita mudanças rápidas e em grande quantidade na história da língua, ao passo que o processo de flexão é muito fechado e se mantém inalterado por longos períodos. (cf. em Inglês *sputnik, beatnik, refusenik; alcoholic, workaholic, radioholic*).

A derivação nem sempre mantém um significado exato, mas pode manter a mesma função (em Inglês, *-al* forma adjetivos de verbos e de nomes: *arrive, dispose: arrival, disposal; brute, option: brutal, optional*). O significado do *-al*, nesses casos, varia: *nasal* significa 'of the nose', *brutal* 'like a brute', *optional* 'is an option' (M. McCarthy, 1991: 318). Em Português, o sufixo *-al* também apresenta o mesmo fenômeno: *cafezal* (plantação de café), *pantanal* (lugar de pântanos) (Rocha, 1998: 111-112)

Derivação própria é a formação de palavras por derivação, que acarreta alguma modificação fonológica, por exemplo, com o acréscimo de elementos novos (prefixos, sufixos).

Derivação imprópria ocorre quando um morfema troca de categoria sem mudar sua forma fonológica (cf. verbo *jantar* e substantivo *jantar*). Conversão é o processo pelo qual palavras mudam de classe sem acrescentar uma afixação específica: *o circular, o brilhante, o ouvinte, o pesar*. Palavras como *mediante, salvo, exceto, apesar (de)* são verbos que se tornaram também preposição. Derivação imprópria é outro nome para conversão.

Derivação parassintética é a formação de uma palavra que contém um prefixo e um sufixo, que não podem ficar de fora da palavra. Assim, *esclarecer* tem o prefixo *es-* mais o radical *clar-* mais o sufixo *-ecer*, (que forma um verbo com certo 'aspecto'), não existindo **clarecer* nem **esclaro* (veja Khedi, 1997: 18). Portanto, não pode ser tirado nenhum dos afixos da formação parassintética: *anoitecer *noitecer *anoite (ficar noite?)*. Alguns verbos parassintéticos apresentam o radical em sua forma arcaica: *abaular*, de *baul*, forma antiga de 'baú'; *esfolar*, de *fole*, na acepção antiga de 'pele'; outros contém uma forma reduzida do radical: *esfulinhar* - de *fulinhagem*; *enjangar* - de *jangada* (Khedi, 1997: 21). Outros exemplos *feliz, infeliz, felizmente, infelizmente - tarde, *tardecer, *entarde, entardecer*. Ocorre frequentemente a derivação parassintética na formação de verbos que exprimem mudança de estado: *engrossar, amadurecer, rejuvenecer*. Veja, ainda adjetivos do tipo *des-alm-a-do*.

Silva & Koch (1987: 33) acham que, na formação parassintética, o que existe é um morfema descontínuo e dizem:

A derivação parassintética distingue-se da anterior (prefixal e sufixal : deslealdade, infelizmente), porque o prefixo e o sufixo são acrescentados a um só tempo ao morfema lexical, constituindo, portanto, um único morfema gramatical, de caráter descontínuo.

Derivação regressiva é o processo de formação de palavras que reduz a forma fonológica da palavra original, gerando uma nova palavra, com mudança de categoria (cf. verbo *pescar* e substantivo *pescar*). O maior número é de substantivos deverbais: *pagar - paga; lutar - luta*. Se a palavra denota ação é derivada de verbo; se não denota ação é o verbo que é derivado do nome. Os deverbais regressivos são extraídos da 1ª ou da 3ª pessoa do singular do presente do indicativo 'daí serem nomes de tema em -o (1ª pessoa) ou

de tema em -a ou -e (3ª pessoa) (Khedi, 1997: 23). Ex: *erro, caça, abate*.

Quando não há a clara presença de uma vogal temática verbal, no final de palavras com derivação regressiva, as vogais que aparecem no lugar da vogal temática verbal são desinências de gênero, como em *embarque, embalo, procura, devassa*. Essas vogais aparecem porque a fonologia não permite sílaba travada por /t, d/, etc. ou *onset* sem vogal. Essas vogais se distinguem da vogal epentética inserida em palavras importadas como *snob - esnobe, club - clube*, etc.

Em Inglês, *hoover* é nome e *to hoover* é verbo. Em Português, *xerox* é nome e *xerocar* é verbo. Malkiel (1978) chama esse fenômeno de derivação interna.

Alguns morfólogos são de opinião que a não existência de certas formas derivadas se deve a um 'bloqueio' causado pela existência de uma outra palavra de mesmo significado. Assim, não existe **ensinador* porque existe a palavra *professor*; não existe **estudador* porque existe a palavra *estudante*. O fenômeno da produtividade nos processos de formação das palavras é uma área que necessita de mais pesquisas.

11. A composição

A composição é o processo de formação de palavras que incorpora dois radicais, em geral, com seus afixos, formando uma única palavra. Fonologicamente, os compostos podem ter autonomia prosódica, isto é, podem ter dois acentos principais, como em *cirurgião-dentista*. Do ponto de vista semântico, os elementos componentes contribuem com parte do significado total do composto. Às vezes, o significado de um é apenas um atributo do significado do outro, como no exemplo acima. Há muitos tipos de "cirurgiões", alguns são dentistas. Outras vezes, o significado do composto refere-se a algo novo, porém, encontrando motivação nos significados individuais das partes, como em *guarda-chuva*. Os compostos são diferentes da simples presença contígua de dois elementos. Assim, quando se diz *azul-marinho*, a palavra *marinho* especifica o tom da cor azul, mas não se torna um advérbio (um adjetivo modificando um outro adjetivo), assim como *cirurgião-dentista* não leva o verbo para o plural, como acontece no caso de o sujeito ser constituído por dois núcleos (Lee, 1997).

Praticamente, podem-se obter palavras compostas juntando qualquer tipo de palavra, e até mesmo sintagmas e frases, como em o

querer-não-querendo-casar e o não-querer-suportando-casar ...
 Nota-se a presença do artigo que, no caso de compostos formados por sintagmas e frases, é sempre masculino. Nos outros casos, a concordância depende dos elementos, se são suscetíveis de fazer concordância e da idéia atribuída aos elementos, a qual, às vezes, pede o singular do atributo, mesmo quando o elemento principal está no plural, como em *escolas-modelo, navios-escola*.

Alguns fonólogos acham que os prefixos formam compostos e não palavras derivadas, como em *supermercado, extraordinário*, etc., tendo em vista a maneira como certas regras fonológicas agem nesses casos. Entre os morfólogos, há uma aceitação maior dos prefixos como elementos da derivação. Nesses casos, se *sobrepujar* é um composto, *submarino*²⁰ também o é. Além disso, alguns pesquisadores interpretam a terminação *-mente* formativa de advérbios como palavra autônoma e os advérbios formados com ela, como palavras compostas e não derivadas. O mesmo vale para as desinências de grau, quando começam com o fonema /z/, como em *mulherzinha, homenzarrão, colherzita, mulherzona*, etc.

12. A estrutura morfológica das palavras

Há duas maneiras de mostrar a estrutura das palavras: uma segmenta todos os morfemas possíveis em todas as formas geradas antes da última. A outra maneira segmenta a partir do radical anterior, o que pode deixar agregado ao núcleo ou raiz alguns afixos, como se fossem partes integrantes e inseparáveis do radical que gera a palavra.²¹

Khedí (1997: 12) faz os seguintes comentários: Na análise de uma palavra como *formalização*, representada {form(a)} + {al} + {iz} + {a} + {ção}, ocorre a derivação de várias palavras, antes de se chegar à forma *formalização*. Segundo ele, o acréscimo de um sufixo depende da 'base' ou do 'radical' ou 'elemento nuclear' ou do 'constituente imediato' anterior. *-al* forma adjetivos de substantivos;

²⁰ É preciso lembrar que não se pode confundir a linguagem escrita ortográfica, que junta esses elementos, com a linguagem oral, base de referência para as análises lingüísticas. A escrita tem suas razões que nem sempre correspondem às unidades lexicais estabelecidas pela Lingüística.

²¹ A opção dos lexicógrafos e a maneira como os dicionaristas organizam o 'seu léxico' têm características próprias: escolhem os adjetivos pelo masculino singular, os substantivos pelo singular, os pronomes pelo masculino singular e os verbos pelo infinitivo.

-izar forma verbo a partir de adjetivos; o sufixo *-ção* só pode se agregar a verbos. Portanto, exige *formalizar + ção*. Para Khedí, *formalização* é um vocábulo derivado do radical *formaliza(r)* mais o sufixo *ção*. O vocábulo *formalizar* é derivado de *formal + izar*. A forma que concatena morfemas não descreve a morfologia da palavra. Uma palavra como *desgostoso* é derivada de *desgosto + oso*, porque o significado é 'cheio de desgosto'. O *des-* é elemento inseparável do radical. Portanto, a palavra não apresenta prefixo. Somente a palavra *desgosto* teria o prefixo *des-*.

A outra abordagem diria apenas que uma palavra como *formalização* é derivada e constituída pelos seguintes morfemas:

form	= base, raiz, núcleo
a	= vogal temática nominal desinência de gênero feminino
Ø	= morfema zero de número = singular
al	= sufixo formativo de adjetivo a partir de substantivo
Ø	= morfema zero de número = singular
iz	= morfema formativo de verbo a partir de adjetivo
a	= vogal temática verbal da 1ª conjugação
ção	= sufixo formativo de substantivo a partir do verbo
Ø	= morfema zero de número = singular

A referência aos morfemas zero é, em geral, dispensada por ser pressuposta, nesses casos. Nota-se que o sufixo *-ção* mantém o gênero original da palavra (*forma*), apesar de terminar em [u].²²

Para os verbos e demais palavras, a orientação dada acima funciona do mesmo modo. Os verbos costumam ser derivados a partir do tema, ou seja, do radical mais a vogal temática verbal, como em *formaliz+a (ção)*.

As palavras podem formar três grupos com relação à sua estrutura: os nomes (substantivos, adjetivos e os pronomes), os verbos e um terceiro grupo formado pelas outras categorias de palavra. Neste último caso, estão as palavras que não têm nenhum tipo de flexão, como as preposições e as conjunções. Os advérbios pertencem a essa terceira categoria, porém, em alguns casos podem ter flexão de grau.

²² É possível supor que a forma antiga fosse [sã ~ sân ou sãm] e que a formação do ditongo [ãu] seja fruto de uma regra morfológica que padronizou esse tipo de terminação na língua. Com isso os sufixos ficaram com a mesma forma das palavras que não apresentam o sufixo, como *pão, chão* (cf. *pam, cham*).

12.1. A estrutura morfológica dos nomes

A estrutura morfológica dos nomes derivados é a seguinte:

{(prefixos)} {raiz} {(sufixos)} {(desinências)}²³

Uma palavra pode ter mais de um prefixo: *des + in + compatibilizar*. Tem apenas uma raiz. Pode ter um ou mais sufixos, que podem ser de qualquer tipo. Alguns desses sufixos trazem consigo alguns elementos das palavras que geraram anteriormente à última forma derivada, como vogais temáticas verbais agregadas a sufixos verbais ou morfema de gênero agregado a formas do participio passado. Em *compatibilização*, ocorre a vogal temática [a] de *compatibilizar*, juntamente com o sufixo *-iz-* formativo de verbos a partir de adjetivos. No exemplo *os perdidos e achados*, encontram-se as vogais temáticas [i], alomorfe de {e} (do verbo *perder*), a vogal temática [a] (do verbo *achar*), o morfema de participio passado [d], a desinência de gênero /u/ e a desinência de número /s/. As formas *ach-* e *perd-* são raízes.

Como uma palavra é derivada de outra, a não ser que seja apenas raiz, o radical conterà a palavra sobre a qual se forma a nova palavra. Assim, como apontou Khedi, a palavra *desgostoso* é formada a partir da palavra *desgosto*. Esta última é o radical de *desgostoso*. A palavra *desgosto* é formada a partir da palavra *gosto*. Portanto, o radical de *desgosto* é a palavra *gosto*. A palavra *gosto* tem o radical *gost* mais a desinência de gênero *-o* e um morfema zero de número.

Para quem considera os morfemas de gênero (ou a última vogal das palavras) como sendo vogais temáticas nominais, a formação morfológica que acaba nelas constituirá o tema nominal da palavra.²⁴

²³ Elementos entre parênteses são opcionais; entre chaves, são morfemas. O termo *raiz* é usado comumente para mostrar a origem histórica ou os morfemas que carregam o significado básico sobre o qual se monta a palavra. O termo *radical* é comumente usado para representar a *raiz* com ou sem outros elementos, na formação de palavras.

²⁴ Um caso curioso e de difícil análise são os pronomes *qualquer, quaisquer* que têm a flexão como infixos. Nota-se nelas a ausência de morfema de gênero e, de certa forma, assemelha-se aos pronomes *tal, tais*. Seria estranho dizer que se trata de uma palavra composta ou de duas palavras. Certamente, não faz sentido dizer que o *quer* que aí aparece seja igual ao *quer* conjunção

Em nenhuma palavra, estará ausente a raiz, podendo, contudo, estarem ausentes os demais tipos de morfemas.

12.2. A estrutura morfológica dos verbos

A estrutura geral dos verbos é a seguinte:

{(prefixos)} {raiz} {(sufixos N/V)} {V. T.} {(desinências verbais)}

Desinências verbais correspondem a: desinência de tempo e modo + desinência de pessoa e número.

No caso dos verbos, entende-se por radical da palavra tudo o que precede a vogal temática. A unidade constituída pelo radical mais a vogal temática chama-se tema. O tema do verbo, portanto, indica a que conjugação ele pertence.²⁵

Em uma palavra como *desrespeitávamos*, o núcleo ou raiz é *respeit*,²⁶ e *des-* é um prefixo de negação. O radical é *desrespeit-a*. A desinência de tempo e modo ('imperfeito do indicativo') é *-va-*. A desinência de pessoa é *-mo-* e a desinência de número é *-s*. Alguns lingüistas interpretam o *-mos* como um morfema com as duas funções: de pessoa e de número.²⁷ No entanto, nos verbos, pode-se

de enunciados como *quer chova quer faça frio, o Governo continua insensível*.

²⁵ Tradicionalmente, as gramáticas indicam as entradas lexicais dos verbos não pelos temas (por exemplo: *cant-a, vend-e, part-i*), mas pelo infinitivo impessoal (por exemplo: *cant-ar, vend-er, part-ir*).

²⁶ Talvez, em Latim, fizesse sentido separar *res* ('coisa') de *pect* ('peito'), mas na Língua Portuguesa, hoje, não faz sentido a segmentação e, portanto, a raiz é *respeit*. Critérios históricos dessa natureza valem somente em se tratando de Lingüística Histórica, não da descrição atual da língua.

²⁷ Um dos motivos pelo qual se optou por agregar os morfemas de pessoa e número em um único morfema, deve-se ao fato de ocorrer um /S/ na segunda pessoa do singular, como em *cant-a-s, vend-e-s, part-e-s*. Na verdade, tal morfema representa somente o morfema de pessoa, uma vez que a forma verbal está no singular e, portanto, não tem nada a ver com o morfema /S/ que representa o plural na primeira e segunda pessoa do plural, como em *cant-a-mo-s, vend-e-mo-s, part-i-mo-s; cant-a-i-s, vend-e-i-s, part-i-i-s*. Achar que o /S/ da primeira e da segunda pessoas do plural não pode ser um morfema distinto de número, somente porque existe um /S/ que ocorre na segunda pessoa do singular, é um argumento sem consistência. Na língua, há muitas ambigüidades. Como foi visto em outra parte, uma palavra como *o dia* tem uma vogal final que é desinência de gênero, como acontece com uma palavra

perfeitamente manter a desinência /S/ como morfema {S} de número plural, como acontece com os nomes. Entretanto, na terceira pessoa do plural, a nasalização /N/ representa os morfemas de pessoa e de número, como em /aʃ + a + va + N/ *achavam* ou /aʃ + a + r + a + N/ *acharão*. Neste último caso, o tepe é o morfema do tempo e modo (infinitivo), o *-a-* que lhe segue é a raiz do verbo *haver* ('hãõ') e o /N/ é a desinência de 3ª pessoa do plural. Os verbos no futuro e no condicional são formados do infinitivo do verbo de significação principal mais o presente do indicativo dos verbos *haver* (futuro) ou do verbo *ir* (condicional).²⁸

Na palavra *repartisseries*, o radical é *repart*, formado de um prefixo *re-* mais um núcleo *part*. O tema é *repart-i*. A flexão de tempo e modo é *-sse-* (imperfeito do subjuntivo); a flexão de pessoa é *-i-* (2ª pessoa) e a flexão verbal de número é *-s* (plural). Na palavra *velejarás*, o radical é *velej* formado da raiz *vel-(a)*, seguida do sufixo *-ej-* formativo de verbos a partir de substantivos.²⁹ Em seguida, vem a vogal temática verbal *-a-* com o tepe representando a flexão de tempo e modo (infinitivo impessoal). Como se trata de um futuro, a última vogal é a raiz do verbo *haver* e o *-s* final é a desinência da flexão de pessoa (2ª). Neste caso, a desinência de número é o morfema zero. O /S/ final, como vimos, em outra abordagem, pode ser interpretado como sendo, ao mesmo tempo, a representação dos morfemas de pessoa e número (2ª pessoa do singular).

A interpretação morfológica da estrutura dos verbos, as conjugações e as derivações dos tempos e modos é um assunto polêmico, apresentando interpretações variadas em alguns casos.

como *a casa*. Não é porque não se pode atribuir ao 'a' final de *dia* a realização do morfema de gênero, que nenhuma outra vogal, como a da palavra *casa* também não pode representar o morfema de gênero. Em Português, a regra geral procura distinguir uma coisa de outra, mas há casos de ambigüidade, nos quais um *-a* final representa um masculino (cf. ainda *sofá, amalgama, etc.*) ou um *-o* final representa um feminino, como em *a concentração, a nação, etc.*

²⁸ Convém lembrar também a forma do imperfeito *havia* e as terminações verbais do mesmo tempo em *-ia*, sendo esse processo de formação do condicional diferente do uso do verbo *ir*. Há muitas opiniões divergentes com relação à conjugação dos verbos em alguns casos, como este.

²⁹ Todavia, há exceção, como a palavra *azulejar* que vem de *azulejo* que, por sua vez, vem de *azul* + o sufixo *-ej-* formativo de verbo. Para os adjetivos, o sufixo mais comum é *-i-*, como em *clarear*. No português arcaico, o som de *j* se derivou do som de *i*, e o mesmo aconteceu com a própria maneira de escrever esses sons.

12.3. A estrutura morfológica das outras palavras

As preposições e as conjunções são categorias de palavras que não têm prefixos, nem sufixos, nem desinências de flexão. Uma palavra como *no* (cf. 'eu fui *no* cinema'), é, na forma subjacente, *em + o*; a preposição propriamente dita é *em* e a vogal *-o* é o artigo masculino singular. Formas como *pro* e *pra* representam as formações subjacentes *para o* e *para a*. No caso da preposição *para*, a vogal final não é flexão nem vogal temática, porque a palavra não pertence à classe dos 'nomes'.³⁰ Nas preposições e nas conjunções, o radical se confunde com a raiz, ou melhor dizendo, como elas não derivam, não existe radical.

13. Morfofonologia

Morfofonologia (ou morfofonêmica) é o estudo da estrutura fonológica dos morfemas. Segundo Trubetskoy (1939), trata-se de um processo em que ocorre mudança de fonemas e não de alofones, como em *elétrico* e *eletricidade* (/k/ ~ /s/). Portanto, todo fenômeno de alomorfa é uma questão de morfofonologia. Em Inglês, as palavras *leaf* e *cuff* fazem o plural com [v] e [f], alofones de dois fonemas diferentes na língua: /f/ e /v/. Portanto, a regra de plural dessas palavras é condicionada morfológicamente. Como diz Dressler (1985a), a morfofonologia intermedia entre a fonologia e a morfologia, sem ser um nível independente na gramática. Alguns fenômenos fonológicos, como sândi externo, dependem sempre da estrutura morfológica de juntura intervocabular ou de frases.

Sândi é o fenômeno de juntar morfemas de tal maneira que ocorra alguma transformação silábica na juntura. Essa transformação pode ser uma ditongação, uma queda de vogal com o deslocamento da vogal seguinte para a sílaba anterior ou o deslocamento de uma consoante da coda para a posição de onset da sílaba seguinte. Quando a transformação acontece dentro de palavras, ocorre 'sândi interno'; quando acontece na juntura de duas palavras, ocorre 'sândi externo'.

³⁰ Historicamente, algumas preposições e conjunções vieram de sintagmas, como *em boa hora* que gerou *embora*, *toda a via* que gerou *todavia*, *com tudo* que gerou *contudo*. Os elementos formativos dessas conjunções foram amalgamados e, hoje, representam apenas uma forma simples, uma raiz morfológica da língua, com a função sintática de conjunção.

Exemplo de sândi interno em Português: *planalto*, palavra composta de *plano* + *alto*; exemplos de sândi externo: *um livro amigo* = *um li.vroa.mi.go*, *casa amarela* = *ca.sa.ma.re.la*, *vem aqui* = *ve.nha.qui*.

Outros fenômenos que apresentam mudanças que são motivadas fonologicamente são apresentados a seguir. Na palavra *casinha*, derivada de *casa* + *inha* ocorre a elisão ou queda (eliminação) da vogal final da primeira palavra. O fenômeno de queda se confunde com o fenômeno de fusão, recebendo o nome especial de crase, quando as duas vogais da fronteira são iguais, como em *tele educação* = *teleducção*. Em vez da vogal final, pode ocorrer o encontro de duas sílabas iguais, fato que a língua procura evitar e para resolver isso, elimina a primeira sílaba desse contexto, como se nota na derivação de palavras como *idolatria* de *ídolo* + *latria*, *Candinha*, de *Cândida* + *inha* e em *bandoso* de *bondade* + *oso*. A esse fenômeno, deram o nome de haplologia. Se as vogais da fronteira forem diferentes e uma delas é uma vogal alta, ocorre uma ditongação e ressilabificação, como em *auto-imagem*.

A troca de *-vel* por *-bil*, quando se acrescenta o sufixo *-idade* a adjetivos é uma fato do sistema atual da língua, que é interpretado como caso de alomorfia e, portanto, uma questão de morfofonologia. Trata-se de uma variação não motivada fonologicamente: seria possível ter *amabilidade*, etc. (com *-zinho*, *-mente* não ocorre *-bil*).

A substituição de /k/ por /s/ mostra também um caso em que a morfologia influencia a fonologia (sistema de fonemas), como em *científico* - *cientificidade*; *presidente* - *presidência*; *autêntico* - *autenticidade*; *prático* - *praticidade*. Há, ainda, fatos como *noz* - *nogueira* em que /s/ é trocado por /g/.

Há um morfema de negação que funciona como prefixo e que sofre alterações fonológicas dependendo dos elementos fonológicos que lhe seguem. Veja os seguintes exemplos: *infeliz*, *inadmissível*, *imaturado*, *iletrado*. Ocorre /iN-/ no primeiro e no segundo exemplos. A realização fonética de ambos depende do segmento seguinte: se for consoante, ocorre a nasalização da vogal anterior com ou sem a presença de uma nasal homorgânica (ou condicionada pela vogal anterior), de acordo com as regras de nasalização da língua. Se o segmento seguinte for uma vogal, ocorre [n]. No terceiro e quarto exemplos, ocorre uma assimilação do arquifonema /N/ com a

consoante seguinte, antes que ocorra a nasalização da vogal anterior, condicionada pelo arquifonema.³¹

Uma questão pouco investigada refere-se à alomorfia dos sufixos. Por exemplo, uma palavra como *variação* é derivada de {var (i) + a + saN + u} e uma palavra como *tensão* é derivada de {teNs + aN + u}. As duas formações derivam substantivos a partir de adjetivos ('vário' - 'tenso') e, portanto, exercem a mesma função morfológica. Porém, o primeiro tem a forma fonológica /saN + u/ e o segundo, /aN + u/. Trata-se, pois, de um caso de alomorfia, mesmo considerando que o /s/ seja uma consoante epentética. Esse tipo de estudo podia se estender a outros casos.

Para quem considera que o /z/ das flexões de grau é um caso de consoante epentética, as variações do tipo *-inho*, *-zinho* são questões de morfofonologia e não apenas de morfologia.

As formações de plural do tipo *lugar* - *lugares*; *animal* - *animais*, ou de gênero, como em *irmão* - *irmã*, ou de ambos, como em *formoso* - *formosa* - *formosos*, bem como em casos como *avô* - *avó*, etc., em que ocorre uma mudança fonológica que remete à palavra como contexto de ocorrência, são casos de alomorfia e, portanto, questões de morfofonologia.

Alguns morfólogos acham que existe um tipo de variação ou de mudança fonológica condicionada pela analogia com outra, como ocorre em *danceteria* (devia ser 'dançateria') por analogia com palavras como *cafeteria* (de 'café').

Sandmam (1991: 54) argumenta que a 'opção de verbos em *-ecer* pelo sufixo *-mento* e não pelo sufixo sinônimo *-ção* é de ordem fonológica: *enfraquecimento* x **enfraquecição*'. Em contrapartida, mostra também (Sandmam, 1992) que o fato de verbos em *-izar* (*flexibilizar*) optarem sempre pelo sufixo *-ção* (*flexibilização* e não **flexibilizamento*) não é de natureza fonológica, mas morfológica. (cf. *lisamento*, *deslizamento*). De fato, as restrições estão corretas, mas o contexto faz sempre referência a elementos morfológicos (os sufixos e a fronteira de morfemas e não somente de sílabas) e, portanto, é uma questão de morfofonologia e não apenas de fonologia. Trata-se de um contexto morfológicamente condicionado.

Como foi visto anteriormente, a representação morfológica pode ser feita através de regras ou de processos. Nessa abordagem, apenas as raízes lexicais são estabelecidas na forma subjacente.

³¹ Embora pareça óbvio, é bom lembrar que não se deve confundir fatos da fala com fatos de escrita. Portanto, está incorreto dizer que o prefixo *des-* tem uma variante *dis-*.

Assim, uma palavra como *desrespeito* fica {Negação} {respeit} {Gênero} {Número}; *desrespeitávamos* fica {Negação} {respeit} {vogal temática verbal 1ª conjugação} {Tempo e Modo} {Pessoa e Número}. Isso significa que muitos elementos aparecem fonologicamente durante a derivação. Como a definição morfológica ficou em uma etapa anterior (num nível superior ou subjacente), a questão da alomorfa não se coloca. Não ocorrendo alomorfa, também não faz mais sentido ter um processo morfofonológico. A noção de processo morfológico acaba com a noção de alomorfa e também com o conceito de morfema zero, uma vez que as entidades morfológicas são definidas como um processo geral e abstrato. Em Inglês, *man + plural* fica *men*; *hit + passado* fica *hit*; *sing + passado* fica *sang*, etc. Essa noção estabelece uma diferença marcante entre o que é um morfema e um morfe e o que eles representam (suas funções).

Para dar conta das funções dos sufixos, por exemplo, que fazem com que uma palavra mude de categoria, a morfologia baseada em processos define os fatos, os quais serão implementados por elementos fonológicos determinados (os afixos). No caso acima, ocorre uma conversão fonológica, ou seja, um processo morfológico que muda a categoria que pertence a uma palavra. Por exemplo, em Inglês, um particípio passado pode ser usado como adjetivo: *broken heart*, um substantivo pode ser usado como verbo: *to Hoover*. Fatos semelhantes ocorrem em Português: *o andar*; *o grilo falante*; *ele tem uma cara conhecida*. O morfema, nesse caso, é uma regra que transforma o particípio passado em adjetivo. A forma verbal e adjetival acabam iguais, mas com funções sintáticas diferentes. A estrutura fonética das palavras é igual, mas, na forma subjacente, a estrutura morfológica é diferente. Veja, ainda, casos como: *um diplomata italiano - um italiano diplomata*.

14. A morfossintaxe

A morfologia mantém relações estreitas com a sintaxe. Muitas palavras precisam ser marcadas no léxico com relação às funções sintáticas que podem exercer. A função da morfologia, nesses casos, é fornecer à sintaxe as palavras adequadas para as funções sintáticas requeridas. O fenômeno poderia ser visto ao contrário: alguns itens lexicais têm marcas especiais no léxico e, portanto, devem ocupar lugares específicos na sintaxe. Essa interface morfologia / sintaxe é

chamada tradicionalmente de morfossintaxe. Apresentam-se, a seguir, alguns fenômenos que marcam os morfemas sintaticamente.

A morfologia derivacional é um dos dois grandes processos morfológicos que marcam os itens lexicais com funções sintáticas específicas, indicando, através de afixos, a que categoria pertence uma palavra. O processo derivacional forma palavras de categorias sintáticas diferentes e acrescenta novos significados.

Morfologia flexional é o processo pelo qual se acrescentam elementos de significado gramatical para exercer funções sintáticas específicas, por exemplo, para marcar a concordância e a regência. Às vezes, um tipo de formação pode gerar palavras que podem ser usadas como de uma categoria diferente da formação original, gerada pela flexão. Isso torna difícil descrever essas formações e usos. Por exemplo, adjetivos podem exercer a função de predicativos; participios e gerúndios podem exercer a função de adjetivos e não de verbos, como em Inglês: *performing seal*, *broken vase*. O fenômeno da flexão é típica de nomes e verbos. Entretanto, em Galês, a preposição tem flexão de pessoa e de número.

Flexão de comparação é um paradigma morfológico, constituído por adjetivos marcados com desinências específicas. Na flexão de comparação, os adjetivos podem ser positivos, comparativos ou superlativos, como acontecia em Latim. Em Inglês, as formas *long*, *longer*, *longest* mostram elementos da flexão comparativa.

Alguns verbos podem ter uma forma flexional para a voz ativa e outra para a voz passiva. Em Latim, as vozes (ativa, passiva e aoristo) apresentam flexões verbais próprias (cf. voz ativa: *amo, amas, amat, amamus, amatis, amant* - voz passiva: *amor, amaris, amatur, amamur, amamini, amantur*). Em Português, a voz passiva é definida pela sintaxe: verbo ser + particípio passado.

Ergativo é uma função sintática de alguns verbos que trocam a função de agente, típica do sujeito, atribuindo-a ao objeto ou a algum outro termo da oração e, ao sujeito, a função de paciente, sem que a frase esteja na voz passiva. Em algumas línguas, há um morfema ergativo específico e não apenas uma marca morfológica no verbo, como acontece com o Português. Veja o exemplo: *o telhado caiu com o vento* e compare-o com *o vento fez cair (derrubou) o telhado*. Absolutivo é o verbo que não é ergativo.

A flexão dos nomes em casos também é uma questão morfossintática. Em Latim, os casos dos nomes eram definidos por declinações de morfemas próprios: nominativo *ros-a*; genitivo *ros-ae*; acusativo *ros-am*, etc. Em Português, os casos são marcados pela posição sintática que a palavra ocupa ou por sintagmas construídos

morfologicamente. Por exemplo, o genitivo é formado pelo sintagma composto de uma preposição 'de' mais um substantivo: 'o livro de Pedro caiu'.

Alguns verbos recebem indicações de funções específicas. Por exemplo, alguns verbos são causativos, outros não. Essas marcas pertencem aos morfemas no léxico e, às vezes, induzem a construções sintáticas especiais. Há muitos 'tipos' de verbos e esses tipos revelam suas funções morfossintáticas.

Há fenômenos morfossintáticos como o foco (ou construção) que é uma marca morfológica que algumas palavras têm para indicar que o objeto direto, o locativo ou o instrumental tornam-se sujeito.

A noção de tempo verbal é uma marca morfossintática dos verbos. Algumas palavras, como os advérbios, também podem significar algo relativo ao tempo. O tempo morfossintático é diferente da idéia de tempo carregado por palavras que não são verbos. O tempo, aqui, é uma idéia gramatical e não lexical. O tempo pode ser presente, passado ou futuro. A formação flexional dos verbos é uma questão típica de morfossintaxe. Além do tempo verbal, há outras funções descritas abaixo.

O modo é outra marca morfossintática dos verbos que caracteriza o significado lexical do verbo como sendo apresentado de modo indicativo, subjuntivo, condicional, optativo ou imperativo.

A modalidade é uma marca morfossintática dos verbos que acrescenta uma atitude do falante em acréscimo ao modo e tempo expresso pelos verbos. A modalidade verbal pode representar uma possibilidade, uma obrigação, uma necessidade, um desejo, polidez, etc.

Aspecto é outra marca morfossintática dos verbos que interpreta-os como revelando algo acabado, completo, imperfeito (cf. os termos 'completivo, perfectivo e imperfectivo ligados à noção de aspecto verbal). É comum tempo e aspecto dependerem de um mesmo sistema de flexão. O aspecto traz uma informação particular a respeito do tempo verbal que se refere, de maneira variada, ao modo como a 'duração desse tempo' é apresentada. Nessa concepção, todo tempo verbal tem alguma referência ao aspecto, razão pela qual alguns autores consideram o aspecto uma forma de flexão verbal (veja comentários de Rocha, 1998, p.225-228).

15. Regras

Regras de estrutura morfológica ou morfemática são aquelas que compõem as formas de base ou formas subjacentes, a partir das quais serão aplicados os processos de derivação até chegar à forma de superfície ou realização fonética.

Regras de ajustamento são aquelas que interferem no processo derivacional, acrescentando, tirando ou mudando de lugar elementos, para evitar estruturas mal-formadas. As regras de epêntese costumam ser regras de ajustamento.

Regras cíclicas são regras cuja aplicação se repete quando encontra contexto favorável, no processo derivacional, dentro do modelo gerativo.

Regras iterativas são regras que se aplicam sobre uma cadeia de elementos (à direita ou à esquerda), modificando o *output* a cada aplicação, terminando quando não há mais contexto para sua aplicação.

No ciclo transformacional, as formas fonológicas de superfície, na fonologia gerativa, apresentam-se indexadas com parênteses que definem contextos (unidades de vários tipos: sintáticas, morfológicas, fonológicas). As regras fonológicas aplicam-se primeiro às seqüências maiores, que não contém os parênteses e, a partir daí, às novas seqüências máximas resultantes (sem parênteses), até que todos os parênteses sejam removidos.

As marcas lexicais são um conjunto de convenções para a interpretação correta de uma matriz lexical. Uma propriedade é marcada ou não marcada com relação a uma regra fonológica. A forma de base não marcada é a raiz e as formas marcadas são os afixos.

Entende-se por Leis Fonéticas o conjunto de leis estabelecidas pelos neogramáticos no séc. XIX, baseadas nas regularidades fonéticas que as línguas derivadas apresentam, a partir das línguas de origem. As exceções são devidas à analogia ou a fatores extra-lingüísticos.

Opacidade é o fenômeno que indica que uma regra fonológica gerativa não encontra suporte na forma fonética, ou seja, a representação fonética não traz nenhuma indicação da existência da regra.

Para-regras ou regras-via (cf. em Inglês *via-rules*) são regras não gerativas que relacionam formas subjacentes, sem introduzir modificações (por exemplo, relacionam *perdiz* com *perdigão*). São regras que não fazem parte do processo de derivação.

Universais lingüísticos são características gramaticais (fonológicas, morfológicas e sintáticas) que são encontradas em todas

as línguas. A idéia de universais lingüísticos é usada também para significar tendências muito comuns e fortes na evolução das línguas. Por exemplo, as línguas tonais tendem a perder os tons e a se tornarem entoacionais, mas o contrário não ocorre.

Matriz é um conceito da fonologia gerativa que interpreta os segmentos em função dos traços distintivos. Como o processo é gerativo, parte-se de uma matriz fonológica, que define os fonemas dos morfemas para uma matriz fonética, que diz como esses morfemas são pronunciados.

Condição de estrutura morfológica ou morfemática é o princípio morfológico que define quais seqüências de morfemas podem ser construídas. Restrições fonológicas também contribuem para definir restrições de combinações de morfemas. Regras fonotáticas são chamadas também de condição de estrutura morfológica.

16. A forma subjacente ou forma de base

As formas de base fazem parte da arquitetura da gramática gerativa. A esse respeito, Silva e Koch (1987: 44) fazem o seguinte comentário, tratando das 'exceções' de gênero:

Na descrição do gênero, como nas descrições lingüísticas em geral, é indispensável delimitar o plano gramatical e o léxico, tendo em vista que a gramática trata dos fatos gerais da língua e o léxico, dos fatos especiais. Assim, a descrição gramatical deve ser completada com as informações de um dicionário ou léxico, que seria constituído, segundo Chomsky (1965), de uma série não ordenada de regras lexicais, englobando todas as propriedades idiossincráticas de cada um dos itens lexicais. Caberia, então, a um dicionário do português, registrar as ocorrências de gênero não explicáveis pelos padrões gerais da gramática.

No modelo da gramática gerativa, o léxico contém as formas de base dos morfemas com todas as marcas gramaticais necessárias para o desempenho de todas as suas funções, em todos os níveis de realização. O léxico pode conter regras abrangentes que se aplicam a conjuntos específicos de itens lexicais.

Quando se diz 'forma de base' o que se quer dizer é a forma fonológica inicial dos morfemas, colocados um ao lado do outro, forma delimitada pela extensão da palavra à qual esses morfemas pertencem, antes que esses morfemas sejam juntados e sejam

aplicados os processos de derivação. Trata-se de uma representação abstrata das unidades morfológicas de uma palavra, que define o estágio mais profundo dos itens lexicais no componente lexical. Essa representação é definida pelas relações morfológicas da língua. A formação da palavra, como unidade, ocorre somente depois que as fronteiras dos morfemas foram removidas. A forma de base é chamada também de representação subjacente ou forma subjacente.

Juntamente com a estrutura segmental dos morfemas, as formas de base apresentam também sua estrutura supra-segmental ou prosódica. Dessa forma, serão especificadas as sílabas, os acentos, etc.

Alguns exemplos de formas de base, delimitadas por barras verticais, sua forma fonética, delimitada por colchetes quadrados, e a forma ortográfica, indicada por itálico, são apresentados a seguir, contendo apenas as informações segmentais:

/para/	[para] <i>para</i> (preposição).
/paN + u/	[pãũ] <i>pão</i>
/paN (i) + fic + a + d + or + a/	[pãnikadora] <i>panificadora</i>
/elstr + ik + idad + i/	[elstrisidãzi] <i>eletricidade</i>
/devagaR + iŋ + u/	[dãivagařĩũ],[dãivagarãĩũ] <i>devagarinho</i> ou <i>devagarzinho</i>

Para uma palavra como *de* (preposição), haverá a indicação de que se trata de uma sílaba átona; porém para uma palavra como *pé*, a indicação dirá que se trata de uma sílaba tônica. Essa informação pode ser obtida através de regras que mencionam a função sintática dessas palavras. Em uma palavra como *estrutura*, na forma de base, a sílaba -*tu*- será marcada como tônica (acento principal) e as demais serão átonas (por convenção geral).

17. Alguns símbolos e contextos usados na morfologia

Marcação de fronteiras: fronteira de palavra #; fronteira de morfema +, fronteira de sílaba \$. As sílabas, às vezes, vêm marcadas com + ou com apenas um ponto (.) ou com um travessão (-).

Parentetização é um modo de isolar morfemas, raízes, radicais, afixos e desinências. Os parênteses mais internos representam a raiz e os mais externos, os afixos. É uma forma de fazer uma distinção entre unidades diferentes contíguas na cadeia da fala. Em lugar dos parênteses, podem ser usados os colchetes.

Indexação (sobrescritos ou subscritos) é uma maneira de rotular os parênteses ou colchetes com os rótulos de categorias morfológicas ou sintáticas.

Barras verticais representam a configuração fonológica de morfemas. O morfema de plural em Português é representado pelo arquifonema /S/ com as barras verticais |S|.

Em lugar das barras verticais, que são usadas mais na morfologia tradicional, a preferência da morfologia gerativa é pela notação através de chaves e não de barras verticais.

Descrição estrutural é o contexto que define as condições necessárias para a aplicação de uma regra gramatical.

Distribuição é o contexto ou lugar em que ocorre uma determinada unidade lingüística. Por exemplo, as vogais têm uma distribuição que se define pela posição de núcleo da sílaba.

Distribuição complementar, como acontece na fonologia, é a configuração do ambiente de ocorrência de elementos, segundo a qual um contexto exclui o outro para uma regra ou para a ocorrência de determinados segmentos. Em outras palavras, trata-se de duas unidades em competição em determinada regra, ocorrendo em contextos diferentes, nunca no mesmo ambiente.

Filtro lingüístico é um processo pelo qual se aplica uma regra, excluindo outras possíveis.

18. A idéia de produtividade em morfologia

Um dos aspectos mais importantes da morfologia é estabelecer os processos de formação de palavras, indicando a possibilidade que a língua tem de criar palavras novas. Ao estudar os morfemas, nota-se que alguns ocorrem com grande freqüência em palavras antigas e em novas formações. Outros aparecem em um conjunto limitado de palavras e, muito raramente, são usados para a formação de uma palavra nova, em geral, sentida pelos falantes como um tanto estranha. A produtividade em morfologia é justamente esse processo que faz com que certas raízes ou prefixos e sufixos sejam usados em muitas palavras antigas e novas. Por exemplo, o morfema *extra* aparece em várias palavras antigas, como *extraordinário*, *extraconjugal*, etc. Recentemente, muitas palavras passaram a ser usadas com esse prefixo, como *extradifícil*, *extraimportante*, *extrafácil*, etc. Como foi dito anteriormente, as terminações latinas que geraram os graus comparativos não são mais produtivas em Português. Embora haja palavras com esse padrão (*melhor*, *menor*,

maior, *pior*, *superior*, *inferior*), não se pode formar outras palavras com ele. A produtividade, neste caso, é zero. Os diminutivos em *-inho* têm uma grande produtividade, como em *carrinho*, *tijolinho*, etc., indo atingir até advérbios, como *devagarinho*. Por outro lado, o diminutivo *-ote* é pouco produtivo: pode-se dizer *caixote*, mas é estranho dizer *casote*, *tijolote* e impossível dizer **devagarote*. É interessante observar a forma *caixotinho* com dois diminutivos, um representado pelo morfema *-ote* e outro, pelo morfema *-inho*. A produtividade do morfema *-ote* é tão baixa que está se tornando parte da raiz, formando uma nova raiz, ou seja, uma palavra específica para 'caixa pequena' - razão pela qual, neste caso, passa a admitir o diminutivo *-inho*.

A produtividade é medida também com relação aos processos morfofonológicos e morfossintáticos. A alomorfa que existe entre oclusiva e fricativa tem uma produtividade alta, e pode ser encontrada em palavras como *ciente* - *ciência*; *presidente* - *presidência*, *contante* - *constancia*, *elétrico* - *eletricidade*, etc.³²

A produtividade pode se aplicar também a outros processos de formação de palavras, como as formas que imitam produtos comerciais, do tipo: *fajutex* (de 'fajuto'), *economex* ('remédio para economia'), etc. Nesses casos, o sufixo *-ex* é altamente produtivo, hoje em dia.

Durante certos períodos da história de uma língua, até os empréstimos podem representar um processo altamente produtivo na formação de palavras novas, sobretudo quando há grandes mudanças culturais, políticas ou tecnológicas. Por volta do século XVI, a Língua Portuguesa recebeu nova leva de palavras latinas, introduzindo um padrão de acentuação proparoxítona que não existia antes (Massini-Cagliari, 1999a). Compare a palavra mais antiga *cadeira* com a posterior *cátedra*. Atualmente, a Língua Portuguesa, como outras, tem incorporado no uso diário inúmeras palavras técnicas, formadas na Língua Inglesa, e adaptadas ao português, sobretudo palavras da área da informática: *fazer back-up*, *fazer download*, etc., ao lado de palavras já mais antigas como *shopping center*, *spa*, etc.³³

³² A terminação *-ância* de palavras como *constância*, *observância*, etc. vem da alofonia que existe entre *ante + ia* e *ância*. Essa formação sufixal já não produz palavras novas em Português, ao que parece.

³³ Por razões de ignorância, de preconceito e de falso nacionalismo, muitas pessoas, consideradas eruditas, gostam de se manifestar na imprensa contra essas palavras novas, trazidas de outras línguas. Esquecem que não são donos da língua e nem têm o controle sobre o que ficará na língua, nem sobre o que não passará de um modismo passageiro. Tal fenômeno sempre esteve

19. Tipologia de línguas pela morfologia

Uma das maneiras de se classificar as línguas do mundo é através da estrutura morfológica das palavras. Uma divisão grande separa as línguas aglutinantes das línguas derivacionais. As línguas aglutinantes simplesmente colocam um morfema ao lado do outro, sem utilizar afixos. As línguas derivacionais são aquelas que apresentam afixos agregados às raízes e aos radicais.

Os lingüistas têm uma divisão mais específica, como apresentada abaixo, salientando que muito raramente uma língua cabe com perfeição absoluta em uma dessas categorias. A classificação é feita pela predominância de certos aspectos. Há outros critérios para se classificar as línguas, como critérios sintáticos: línguas tipo S-V-O (sujeito - verbo - objeto), S-O-V (sujeito - objeto - verbo), línguas do tipo O-S-V (objeto - sujeito - verbo), etc. Há línguas do tipo tópico / comentário e línguas do tipo sujeito / predicado, etc. (Greenberg, 1974; Comrie, 1989).

As línguas isolantes são consideradas 'analíticas' e as aglutinantes e flexivas, de 'sintéticas'.

Uma língua isolante é aquela em que há uma correspondência bem próxima entre morfemas e palavras, como o Chinês e o Vietnamita. Nas línguas isolantes, todas as palavras são invariáveis.

Aglutinante é o tipo de língua que se serve de morfemas como 'afixos invariantes' – com o objetivo de representar fatos gramaticais - para formar palavras, podendo ocorrer vários morfemas em uma única palavra. Nessas línguas, as fronteiras dos morfemas estão sempre claras e não costumam sofrer alterações (por exemplo, processos de sândi), uma vez que cada morfema tem uma forma praticamente invariante. Exemplos de língua desse tipo são o Japonês e o Turco.

Língua flexiva é aquela que apresenta conjugações e declinações. Nessas línguas, as funções morfológicas podem não ser claras à primeira vista e nem todos os morfemas raízes podem ser usados isoladamente. O Grego e o Latim são línguas flexionais.

É chamada de polissintética a língua cujas palavras são formadas por muitos afixos, de todos os tipos. As línguas Esquimós são desse tipo.

presente na vida de todas as línguas. A Língua Inglesa, por exemplo, tem a maioria de suas palavras oriundas do Latim, em geral, através do Francês.

20. Conclusão

Esta segunda parte do livro apresentou de modo sucinto noções gerais de morfologia. É importante que o leitor se familiarize com essas noções, pois delas depende a compreensão de muitas análises apresentadas a seguir.

Como se pode notar, a morfologia é uma parte da gramática que tem longa tradição de estudos e, por essa razão, juntou idéias, conceitos e definições dentro de quadros teóricos muito diferentes, como o da gramática normativa tradicional, do estruturalismo, da gramática funcional e da gramática gerativa moderna. Como não é o propósito desta obra distinguir e discutir dentro de uma única teoria, foi necessário apresentar as noções básicas de morfologia em uma forma enciclopédica. Alguns conceitos básicos foram apenas definidos, outros mereceram algumas explicações mais detalhadas. Apesar do caráter um tanto eclético das idéias, houve um esforço para traduzir, numa visão lingüística moderna, o que veio de outros tempos.

O quadro teórico da teoria da otimalidade e suas implicações morfológicas não foram colocados, estando reservados para eles uma seção especial no livro.