

MORFOLOGIA DISTRIBUÍDA: ORIGEM E MOTIVAÇÕES

João Paulo Lazzarini Cyrino

Conceitos-chave do Capítulo:

- O estudo da morfologia na Teoria Gerativa
- A Hipótese Lexicalista
- O modelo da Morfologia Distribuída
- A interpretação não-lexicalista de *Remarks on Nominalizations*

Com o artigo de Morris Halle e Alec Marantz, *Distributed Morphology and the Pieces of Inflection*, publicado no início dos anos 1990, surge a Morfologia Distribuída oferecendo uma alternativa à visão de arquitetura da Gramática que vinha sendo largamente assumida até então: a Hipótese Lexicalista.

A Hipótese Lexicalista é caracterizada por uma arquitetura da Gramática com dois componentes gerativos: léxico, basicamente responsável pela formação de palavras, e sintaxe, basicamente responsável pela formação de sentenças. Essa arquitetura surgiu dentro da Teoria Gerativa atribuída a um desenvolvimento da proposta de Chomsky, em 1970, com *Remarks on Nominalization*, e teve um impacto muito grande no que se compreendia e, em alguma medida, se compreende hoje, a respeito da relação entre morfologia e sintaxe.

A postulação de dois componentes gerativos na Gramática, porém, traz consigo o potencial de se criar um abismo entre os princípios de formação de palavras e os princípios de formação de sentenças. Tal abismo foi frequentemente questionado ao longo

do desenvolvimento da Teoria Gerativa e uma grande discussão dentro da Hipótese Lexicalista é justamente até que ponto palavras são formadas somente no Léxico: há fenômenos, como a flexão verbal, que sugerem que alguns aspectos da morfologia dependem de princípios sintáticos. De fato, grande parte dos trabalhos que consideram a Hipótese Lexicalista, consideram-na em uma versão que assume que ao menos parte dos fenômenos de morfologia flexional se dá no componente sintaxe.

A Morfologia Distribuída, em contraste, propõe que palavras e sentenças sejam derivadas em um único componente gerativo, a sintaxe. Dessa forma, a relação entre morfologia e sintaxe tende a ser mais transparente, com os mesmos princípios regendo a formação de palavras e de sentenças, e o modelo tende a ser mais econômico. Embora esse tipo de concepção já tenha sido empreendido em diferentes momentos da história da Teoria Gerativa, a Morfologia Distribuída é consequência de um longo percurso de discussões que remontam ao surgimento da própria Hipótese Lexicalista.

Neste capítulo inicial, abordamos em detalhe esse histórico do estatuto da formação de palavras em Teoria Gerativa, que culminou na proposta da Morfologia Distribuída. Iniciamos o percurso primeiramente explorando as origens e desenvolvimento da Hipótese Lexicalista. Em seguida, apresentamos o modelo da Morfologia Distribuída e algumas das características que o embasam. Por fim, discutimos uma interessante controvérsia sobre o texto tido como fundacional da Hipótese Lexicalista, *Remarks on nominalization* (CHOMSKY, 1970).

1. Origens e desenvolvimento da Hipótese Lexicalista

Embora a Hipótese Lexicalista não seja representada por uma única concepção de Gramática, entre as abordagens que

partem dela é comum a ideia de que a derivação de palavras segue princípios distintos dos da derivação de sentenças. Essa ideia pode ser assumida de forma mais ou menos radical, mas sempre envolverá conceber uma Gramática em que atuam dois componentes gerativos: a sintaxe, responsável pela formação de sentenças e o léxico, responsável em um grau maior ou menor pela formação de palavras. Embora a formulação da Gramática com dois componentes gerativos possa ser considerada cientificamente custosa, há muito tempo ela tem sido assumida em grande parte dos trabalhos produzidos em Teoria Gerativa. Além disso, ela é responsável por um retorno da concepção da morfologia enquanto um nível linguístico com princípios e mecanismos próprios, algo que havia sido paulatinamente abandonado na segunda metade do século XX.

Nesta seção, abordamos as origens da Hipótese Lexicalista e como ela se consolidou como um padrão na formulação de muito do que se entende até hoje como arquitetura da Gramática. Para isso, tecemos, primeiramente, algumas considerações sobre como a morfologia enquanto nível linguístico vinha sendo encarada desde os inícios do século XX até sua metade, com o surgimento da Teoria Gerativa.

1.1 A autonomia da morfologia

O estudo da estrutura e da formação de palavras era de grande importância nos trabalhos em linguística estrutural, na primeira metade do século XX. Em 1921, no manual *Language: gente Introduction to the Study of Speech*, Edward Sapir, por exemplo, tipificou as línguas com base em como suas palavras se estruturavam. Isso aprimorou as classificações herdadas do século XIX – *línguas analíticas, aglutinantes, sintéticas, polissintéticas*. Nesse sentido,

a tipologia morfológica das línguas relaciona-se não somente à complexidade morfológica das palavras, mas também aos processos gramaticais que criam essa complexidade (composição, afixação, reduplicação, etc.) e aos conceitos veiculados nesses processos. Nos anos 1940 e início dos 1950, verifica-se também um movimento de investigação – representado principalmente por Zellig Harris, Bernard Bloch e Eugene Nida – a respeito do estatuto do morfema, definindo-o ou problematizando-o. Nesse sentido, pode-se dizer que até esse momento, a morfologia era considerada uma área autônoma, com questões específicas sobre si própria.

No entanto, nas décadas de 1950 e 1960, quando surge a Teoria Gerativa, a preocupação com questões específicas de morfologia sofre um declínio. Com a Sintaxe se estabelecendo como um campo de investigação, questões de ordenamento dos morfemas rapidamente passaram a ser abordadas no mesmo domínio de investigação que questões de formação de sentenças. A publicação do *The Sound Pattern of English* por Noam Chomsky e Morris Halle em 1968 também consolida uma visão bastante abrangente de fonologia que permite que questões sobre alomorfia sejam englobadas nesse domínio de investigação.

Enquanto essa absorção da morfologia por outros níveis linguísticos pôde ser compreendida como um avanço no sentido de economia de princípios, há certamente consequências negativas para esse movimento teórico. Ao longo dos anos 1960, houve uma crescente percepção de que alguns fenômenos relacionados à formação de palavras estão sujeitos a idiosincrasias que a sintaxe ou a fonologia não poderiam acomodar de forma elegante. Concretamente, o estudo das nominalizações lançado por Chomsky (1970) é comumente responsabilizado por trazer a morfologia de volta a uma posição de autonomia. Esse momento, que culmina com a proposta da Hipótese Lexicalista, é visto por alguns morfólogos como o renascimento da morfologia (ARONOFF, 1976).

1.2 As nominalizações e a origem da Hipótese Lexicalista

A origem da Hipótese Lexicalista é geralmente atribuída a um texto de Chomsky publicado em 1970, intitulado *Remarks on nominalization*. Na época, os gerativistas desenvolviam seus trabalhos dentro de um modelo denominado Teoria Padrão, originário do livro *Aspects of the Theory of Syntax*, de 1965, também de autoria de Chomsky.

O raciocínio básico desse modelo era o de que havia uma lista de entradas lexicais (palavras) já previamente categorizadas como nomes (substantivos e adjetivos), verbos, preposições etc. Esses itens lexicais se combinavam em estruturas sintáticas constituindo uma Estrutura Profunda. Essa Estrutura Profunda estaria sujeita a regras transformacionais, que a modificavam de forma restrita, sem alterar o significado final. Essas transformações poderiam envolver tanto alterações de ordem sintática, como mudanças morfológicas. O modelo era, portanto, dividido em dois componentes principais: o componente base, responsável pela geração da Estrutura Profunda, e o componente transformacional, responsável pelas regras transformacionais.

Um típico exemplo de regra transformacional seria a passivização:

- (1) a. João comprou um livro.
- b. Um livro foi comprado por João

Segundo a visão que se tinha naquele momento, a única diferença entre as sentenças (1)a e (1)b acima estava na estrutura sintática. Enquanto uma apresentava um núcleo verbal V, um argumento A na posição de sujeito e um argumento B na posição

de objeto, a segunda apresentava um núcleo V participial, com o argumento B na posição de sujeito e o argumento A introduzido por preposição. No modelo transformacional, a sentença (1)b seria derivada da sentença (1)a a partir de uma regra denominada passivização, que participializa (*part*) o núcleo verbal V e altera o posicionamento de seus argumentos A e B, com uma preposição (*prep*) introduzindo o argumento A (ex. (2)).

- (2) a. A...V... B
 b. B...V(*part*)...*prep* A Regra transformacional de passivização

As diferenças entre as línguas estavam associadas, principalmente, ao formato e aplicação dessas regras transformacionais e boa parte dos trabalhos em Teoria Gerativa consistia em propor regras transformacionais que pudessem derivar os fenômenos sintáticos e morfológicos vistos nas línguas.

Em *Remarks on nominalization*, por outro lado, encontra-se uma reflexão relevante a respeito das limitações das regras transformacionais. O texto problematiza a ideia que vinha sendo defendida a partir de Lees (1960) de que nominalizações – derivações de substantivos a partir de verbos – seriam resultados de uma regra transformacional.

É possível estabelecer uma comparação entre a nominalização e a passivização, para a qual já havia sido proposta uma regra transformacional (ex. (2)). Tomando os dados a seguir, vemos que a versão nominalizada dos dados em (3), em (4), e sua versão passivizada em (5), além de corresponderem semanticamente entre si, podem ser derivadas a partir de regras semelhantes. Enquanto passivizações seriam derivadas pelo modelo visto em (2) b, as nominalizações poderiam ser derivadas por um modelo como V(nominalizado) sto B...*prep* A:

- (3) *Sentença sem transformação: A...V...B*
- a. João recusou a proposta.
 - b. João destruiu a cozinha.
- (4) *Sentença nominalizada: V(nominalizado)...prep B...prep A*
- a. A recusa da proposta por João.
 - b. A destruição da cozinha por João.
- (5) *Sentença passivizada: B...V(part)...prep A*
- a. A proposta foi recusada por João.
 - b. A cozinha foi destruída por João.

No entanto, há um grande problema em dizer que sentenças nominalizadas são resultados de regras transformacionais assim como as passivas: ao menos as nominalizações derivadas (baseadas no particípio) são imprevisíveis em forma e significado. Por exemplo, é difícil prever qual morfema nominaliza cada tipo de verbo: *destruir* pode se tornar um substantivo com a adição de *-ção*, resultando em *destruição*, mas o mesmo não ocorre com *esquecer*, que resulta em *esquecimento* após o processo de nominalização (*esqueceção* não é uma forma válida na língua).

Outro fenômeno curioso é o de que a semântica dos substantivos derivados de verbos pode apresentar variações a princípio difíceis de serem descritas por uma regra transformacional, que deve a todo custo manter o paralelismo semântico. Assim, *ensinamento* é entendido como “aquilo que é ensinado”, mas *tratamento* jamais é entendido como “aquilo que é tratado”, e sim como o ato ou método de tratar.

Tendo em vista dificuldades desse tipo, Chomsky conclui que, para as nominalizações derivadas serem abordadas como resultado da aplicação de uma regra transformacional, seria necessário um sistema que estipulasse que formas verbais devessem poder se

nominalizar com diferentes características semânticas entre si. Essa solução, contudo, violaria a essência do que seria uma regra transformacional: uma regra aplicável à estrutura sintática e não à semântica dos elementos envolvidos.

Chomsky, então, apresenta como solução algo que ele mesmo denomina *posição lexicalista*. Tratava-se de “estender as regras da base para acomodar os substantivos nominalizados diretamente”. Dessa solução, há um entendimento quase geral de que Chomsky apontava para a possibilidade de que o componente lexical do modelo, que ora era apenas uma lista, se tornasse um componente derivacional com suas próprias transformações.

Dizemos que esse entendimento é quase geral porque existe uma importante controvérsia levantada posteriormente pelo linguista Alec Marantz a respeito do que Chomsky teria realmente proposto em seu texto. Essa controvérsia será discutida em uma seção específica ao final deste capítulo. Por ora, cabe dizer que *Remarks on nominalization* é tido como um marco na Teoria Gerativa por apontar, em tese, que nem todos os fatos derivacionais das línguas naturais se dão por meio de regras transformacionais, mas que deve haver ao menos alguns tipos de derivações realizadas em um componente lexical.

1.3 A Morfologia Lexical

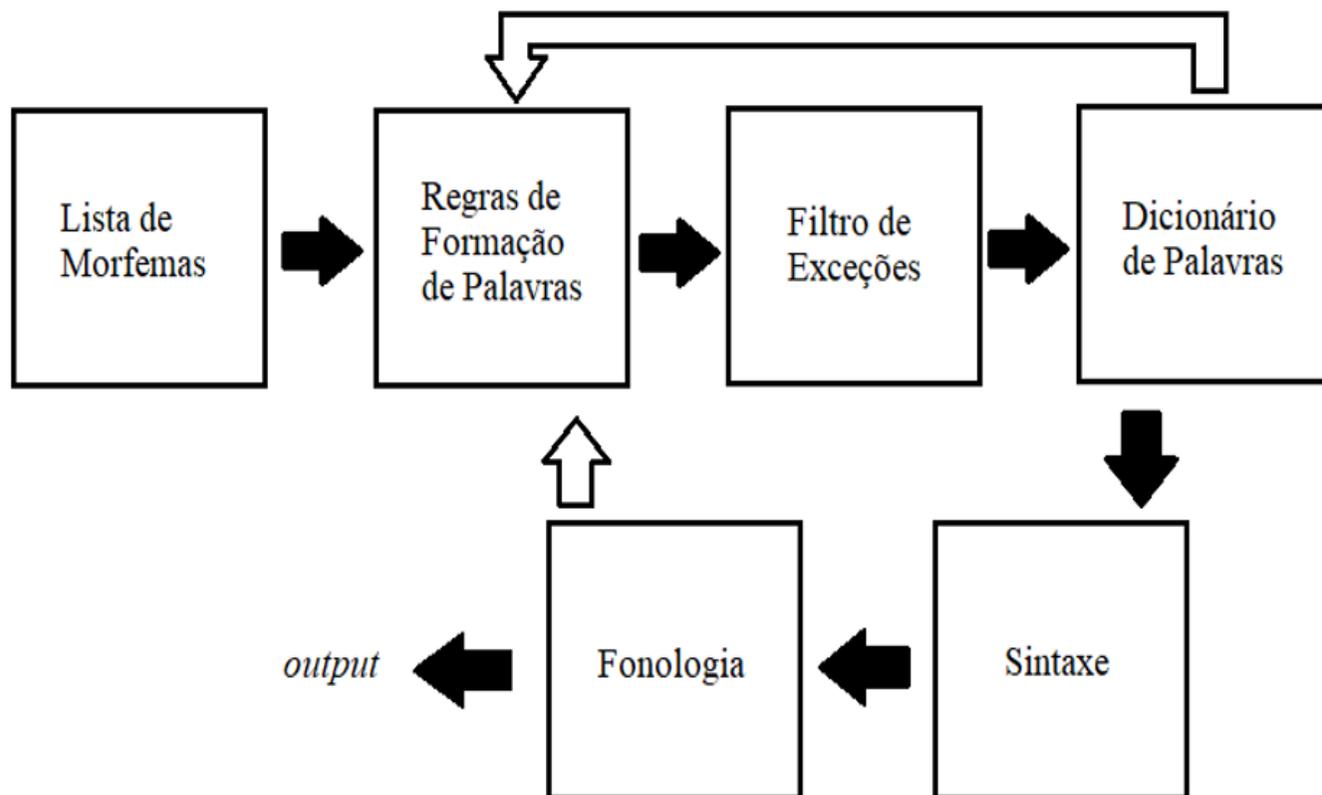
Chomsky não desenvolve um modelo lexicalista de forma detalhada em seu texto de 1970, embora evidencie que ao menos algumas palavras complexas são mais bem explicadas como derivações lexicais do que como resultados de uma transformação sintática.

O primeiro a chamar a atenção para a necessidade de se estudar com mais seriedade e independência os processos de

formação de palavras é Morris Halle, em 1973. Em seu texto *Prolegomena to a Theory of Word Formation*, o autor chama a atenção para a necessidade de se representar formalmente o conhecimento do falante nativo a respeito da formação de palavras.

Halle esboça um modelo com o seguinte formato:

Figura 1 – O Modelo de Halle (1973).



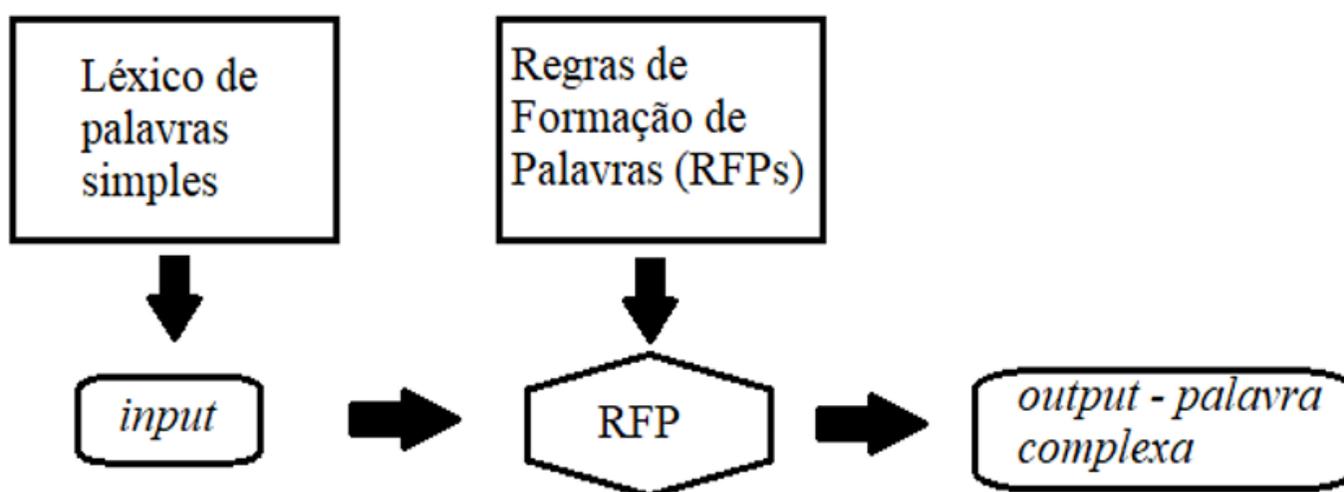
Fonte: traduzido e adaptado de Halle (1973, p. 8).

A Lista de Morfemas a partir da qual tem início o processo de formação de palavras inclui afixos e raízes. Os elementos dessa lista são combinados por meio de Regras de Formação de Palavras. Os resultados dessas combinações estão sujeitos a um Filtro de Exceções e podem ser alterados por meio de uma intervenção *ad hoc*. Os resultados de todo esse mecanismo estão representados no Dicionário de Palavras, que, por sua vez, alimenta a derivação sintática.

É importante considerar dois fatores a respeito do que Halle inclui em seu texto. Por um lado, é possível reconhecer Halle como mais radical do que Chomsky, ao considerar que toda a formação de palavras deve se dar fora da sintaxe. Por outro, o modelo de Halle não é proposto com todo o rigor de um modelo teórico, sendo mais uma forma de instigar um debate e abrir um campo de investigação, do que a criação de uma proposta lexicalista propriamente.

Como consequência das provocações lançadas por Halle, pode se mencionar o trabalho de Aronoff (1976), *Word Formation in Generative Grammar*. A proposta de Aronoff se destaca por explicar a formação de palavras com base em regras que criam novas palavras a partir de outras, e não com base em uma lista de morfemas. Nesse sentido, o Léxico de Aronoff é constituído de um léxico de palavras simples e de uma lista de regras que formam palavras complexas a partir de palavras mais simples. Afixos não apresentam um estatuto próprio nessa teoria e são vistos como parte do *output* de uma regra de formação de palavras.

Figura 2 – Ilustração do Léxico de Aronoff (1976).



Fonte: elaboração própria.

É comumente considerado que a ideia do léxico enquanto lugar da morfologia se desenvolve a partir de Aronoff (1976) e

diversas outras propostas que consideram que palavras e sentenças são formadas em lugares distintos surgem a partir desse momento. Dentre essas, é relevante mencionar Pesetsky (1979) e Kiparsky (1982a), com a Morfologia (e Fonologia) Lexical, que traz forte influência da Fonologia para a compreensão dos processos de formação de palavras. Esse tipo de modelo será abordado no capítulo *Domínios de localidade na forma fonológica*.

1.4 A Hipótese Lexicalista para sintaticistas

Remarks on nominalization não tem consequência apenas entre os morfólogos. Dentro dos estudos em Sintaxe era possível observar um crescente número de trabalhos que serviam de base para explicar fenômenos morfossintáticos com base nas diferenças entre a derivação feita no léxico e na sintaxe. Os trabalhos de Jackendoff (1975) e Wasow (1977) são exemplos disso, mostrando as restrições semânticas a que cada componente estaria sujeito. Pode-se dizer que a Hipótese Lexicalista também encontra origem nesses trabalhos.

Nos estudos em teoria sintática é comum o uso de termos como Hipótese Lexicalista Forte, que remete a todo tipo de morfologia como derivada no léxico, bem como Hipótese Lexicalista Fraca, que, por sua vez, considera que morfologia flexional é derivada na sintaxe e morfologia derivacional, no léxico. Essas diferentes abordagens trazem grandes diferenças na análise de fenômenos sintáticos e são amplamente discutidas nos anos 1980, sendo a Hipótese Lexicalista Fraca comumente mais aceita entre os sintaticistas.

De fato, para a maior parte dos estudos em Sintaxe Gerativa a partir dos anos 1980, a Hipótese Lexicalista Fraca é um pressuposto, estando presente nas formulações iniciais do modelo de Regência

e Ligação (CHOMSKY, 1981, 1986^a) e do Programa Minimalista (CHOMSKY, 1993, 1995).

O tratamento de flexão e da derivação no modelo da Morfologia Distribuída será abordado no Capítulo 7.

2. Morfologia Distribuída

Com o que vimos na seção anterior é possível mencionar duas características que tornam a Hipótese Lexicalista atrativa. A primeira é que a existência do léxico soluciona as imprevisibilidades que fenômenos morfológicos podem apresentar, sem complicar o funcionamento das transformações sintáticas. A segunda é que a existência do léxico dá maior autonomia para a morfologia, que – ao menos na versão forte da Hipótese Lexicalista – não se vê absorvida na sintaxe ou na fonologia. A existência de dois componentes gerativos – léxico e sintaxe – é uma fragilidade, no entanto, podendo ser questionada quando se conseguisse explicar os fenômenos com apenas um deles.

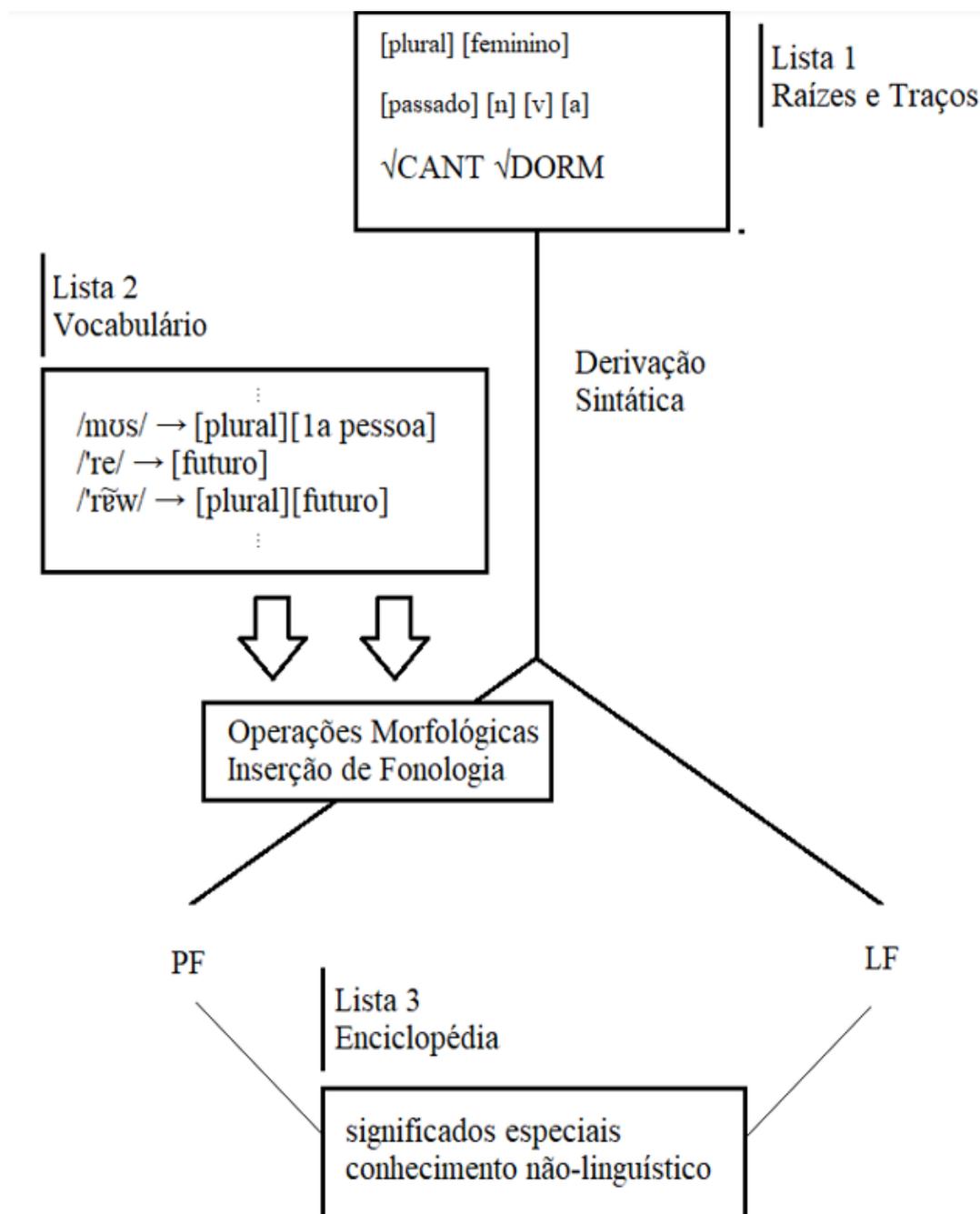
O início dos anos 1990 na Teoria Gerativa é marcado por discussões a respeito de elegância e adequação explicativa dos modelos teóricos. Buscavam-se teorias econômicas, com poucos princípios, que explicassem uma grande quantidade de fenômenos. No contexto dessas discussões, que deram origem ao Programa Minimalista, surge a Morfologia Distribuída: um modelo de Gramática que rompe com a Hipótese Lexicalista.

Morris Halle e Alec Marantz publicam, em 1993, o artigo intitulado *Distributed Morphology and the Pieces of Inflection*, o texto fundacional da Morfologia Distribuída. No novo modelo, as informações que estavam contidas nas entradas lexicais de um componente único, o léxico, passam a se encontrar distribuídas por todas as etapas da derivação sintática. A formação de palavras deixa de se confinar, portanto, em um reduto com princípios e

propriedades próprias.

A proposta pode se adequar tanto a uma arquitetura com base no modelo de Regência e Ligação, com Estrutura Profunda, Superficial, Forma Lógica e Forma Fonética, como a uma arquitetura com base no que tem sido proposto no Programa Minimalista, sem os níveis de representação intermediários (Estrutura Profunda e Superficial). O mais importante, no entanto, é que, em vez de um léxico gerativo que alimenta a derivação sintática, há três Listas que interagem com essa derivação em momentos distintos. A figura a seguir esquematiza essa arquitetura:

Figura 3 – arquitetura da Morfologia Distribuída.



Fonte: elaboração própria.

Os elementos primários da teoria são traços gramaticais abstratos, tais como [plural], [feminino], [n], [futuro] etc., e raízes – elementos sem nenhuma especificação com relação à sua categoria gramatical. Esses traços e raízes compõem uma primeira lista, conhecida como Lista 1, que será detalhada no capítulo *Lista 1: traços morfossintáticos e raízes*.

Os elementos da Lista 1 adentram a derivação sintática como

feixes de traços abstratos, ainda sem nenhum correspondente fonológico. A sintaxe combina esses feixes de traços e raízes e, em um determinado momento, a derivação sofre uma bifurcação, chamada de *spell-out*: uma cópia do que tenha sido derivado até esse momento segue para a interface com o sistema Sensorio-Motor como Forma Fonética (PF), e outra cópia segue para interface com o sistema Conceitual-Intencional como Forma Lógica (LF).

Essa bifurcação não é novidade da Morfologia Distribuída, mas tem seu lugar em grande parte dos modelos gerativistas. A novidade da Morfologia Distribuída é que, no momento da bifurcação, a sintaxe ainda está operando com material abstrato, sem correspondência fonológica. Essa correspondência fonológica só ocorre no caminho para Forma Fonética, quando entra em cena a Lista 2, a ser apresentada em detalhes no capítulo *Lista 2: O Vocabulário*.

A Lista 2 é conhecida por carregar os Itens de Vocabulário. Esses Itens de Vocabulário nada mais são do que regras de correspondência entre os nós terminais abstratos, derivados na sintaxe a partir da combinação de feixes de traços da Lista 1, e feixes de traços fonológicos. Um exemplo de Item de Vocabulário vem a seguir:

- (6) *Concordância de primeira pessoa do plural em turco:*
 /Iz/ → [1ª pessoa],[plural] / v__

A regra acima pode ser lida como: insira a forma fonológica /Iz/ onde forem encontrados os traços de primeira pessoa e plural ([1ª pessoa], [plural]) em um ambiente verbal (v__). Para entender exatamente o que essa regra representa, tomemos alguns dados do turco:

- (7) a. **geliyorum** *eu estou vindo* e. **geliyorus** *nós estamos vindo*
 b. **gelirim** *eu venho* f. **geliriz** *nós vimos*
 c. **geleceğim** *eu virei* g. **geleceğim** *nós viremos*
 d. **geldim** *eu vim* h. **geldik** *nós viemos*

Acima, vemos o verbo *gelmek* (vir) em turco, conjugado em primeira pessoa do singular (7)a-d e do plural (7)e-h nos tempos presente contínuo, presente simples, futuro e pretérito. Esses tempos são expressos pelos seguintes morfemas respectivamente: -Iyor, -Ir, -AcAk, -d.

Alguns esclarecimentos sobre a fonologia devem ser feitos. Primeiramente, os fonemas I e A são arquifonemas, representando realizações possíveis de acordo com o fenômeno da harmonia vocálica na língua: a depender da vogal anterior, I pode se realizar como -i, -u, ü ou -ı e A pode se realizar como -a ou -e. Outro esclarecimento, a consoante /k/ cai facilmente em contextos intervocálicos, sendo essa queda representada pela letra ğ na ortografia turca.

Feitos esses esclarecimentos, é possível observar que o Item de Vocabulário em (6) é o responsável por inserir o expoente /Iz/ em (7)e-g, que na articulação fonética será realizado respectivamente como [uz], [iz] e [iz]. Esse processo em que um nó terminal sintático ganha uma correspondência fonológica é denominado Inserção de Vocabulário. Trata-se de um mecanismo central para a Morfologia Distribuída e que segue alguns princípios cruciais. Um deles é importante para continuarmos entendendo o que sucede nos dados em (7): um feixe de traços abstratos pode ser associado a qualquer Item de Vocabulário especificado para um subconjunto dos traços desse mesmo feixe.

Para ilustrar, temos a seguir uma representação bastante esquemática de como se encontrariam na derivação sintática, ainda sem correspondência fonológica, os feixes de traços para a concordância para o presente contínuo, o presente simples, o futuro

e o pretérito do verbo *gelmek* (vir) na primeira pessoa do plural:

- (8) a. {[1ª pessoa], [plural], [presente], [contínuo]}
 b. {[1ª pessoa], [plural], [presente]}
 c. {[1ª pessoa], [plural], [futuro]}
 d. {[1ª pessoa], [plural], [pretérito]}

De acordo com o que foi apontado, o Item de Vocabulário em (6) poderia inserir a forma /Iz/ em cada uma das estruturas vistas acima. Isso porque os traços [1ª pessoa] e [plural] presentes no Item de Vocabulário são um subconjunto dos traços vistos nas estruturas em (8).

Contudo, a regra derivaria uma forma de pretérito inexistente em turco. A forma **geldiz* entraria em (8)d e, como vemos em (7) h, a forma esperada nesse caso é *geldik*. Essa inconsistência serve para explicar outra propriedade crucial da Inserção de Vocabulário: Itens de Vocabulário competem entre si para a inserção de fonologia em um feixe de traços sintáticos. Essa competição é vencida pelo Item de Vocabulário que esteja especificado para um conjunto de traços maior dentro daquele feixe. Considerando essa propriedade, é possível explicar a distribuição das formas da primeira pessoa do plural nos tempos verbais do turco com dois itens de vocabulário:

- (9) a. /Iz/ [1ª pessoa],[plural] / v____
 b. /Ik/ → [1ª pessoa],[plural],[pretérito] / v____

Com a existência de um Item de Vocabulário especificado para o tempo [pretérito], ocorre que, enquanto a forma /Iz/ pode, em tese, ocorrer nos quatro tempos, a forma /Ik/ ganha a competição para o futuro em (8)d, pois ela corresponde a um Item de Vocabulário mais específico para os traços presentes naquela estrutura do que a forma /Iz/. A seguir podemos ver os resultados da aplicação desse

mecanismo.

- (10) a. {[1ª pessoa], [plural], [presente], [contínuo]} → geliyoruz
 b. {[1ª pessoa], [plural], [presente]} → geliriz
 c. {[1ª pessoa], [plural], [futuro]} → geleceğiz
 d. {[1ª pessoa], [plural], [pretérito]} → geldik

Interessantemente, se adicionarmos mais um Item de Vocabulário, apenas especificado para o traço [1ª pessoa], conseguimos explicar toda a distribuição das primeiras pessoas do singular e do plural vistas em (7):

- (11) a. /Iz/ → [1ª pessoa],[plural] / v____
 b. /Ik/ → [1ª pessoa],[plural],[pretérito] / v____
 c. /Im/ → [1ª pessoa] / v____

Todos os dados de primeira pessoa do singular em turco ocorrem com a mesma marca de concordância *-Im* (exs. (7)a-d), não sendo necessário nenhum outro item específico para tempo. Além disso, como a forma de plural é mais específica, não é necessário incluir um traço [singular].

- (12) a. {[1ª pessoa], [presente], [contínuo]} → geliyorum
 b. {[1ª pessoa], [presente]} → gelirim
 c. {[1ª pessoa], [futuro]} → geleceğim
 d. {[1ª pessoa], [pretérito]} → geldim
 e. {[1ª pessoa], [plural], [presente], [contínuo]} → geliyoruz
 f. {[1ª pessoa], [plural], [presente]} → gelirsiniz
 g. {[1ª pessoa], [plural], [futuro]} → geleceğiniz
 h. {[1ª pessoa], [plural], [pretérito]} → geldik

A Morfologia Distribuída também prevê que algumas

operações possam atuar imediatamente antes ou imediatamente depois da inserção de vocabulário, o que permite uma série de resultados interessantes. O capítulo *Operações morfológicas* descreverá esses mecanismos com mais precisão. Por ora, cabe lembrar que tudo o que é realizado em torno da inserção de vocabulário terá impacto somente na realização fonológica da estrutura, e não em seu significado.

Há ainda uma Lista 3, denominada Enciclopédia, que apresentaremos no capítulo *Lista 3: Enciclopédia*. Essa lista tem por função armazenar significados especiais, tais como expressões idiomáticas. Como a Morfologia Distribuída não opera com palavras, e sim com feixes de traços abstratos, a Enciclopédia não tem seu escopo limitado apenas a palavras, podendo listar significados especiais para unidades que variam desde morfemas até frases inteiras. Isso é especialmente relevante para dar conta de diferentes níveis de expressão idiomática: desde afixos, por exemplo, diminutivos como marcas depreciativas (*livrinho*, livro de qualidade desprezível vs. livro pequeno), passando por palavras com significados especiais (*gato*, animal vs. homem bonito vs. instalação elétrica irregular) até expressões idiomáticas envolvendo verbo e objeto e um sujeito variável (*X chutou o balde*) e expressões idiomáticas envolvendo uma sentença inteira (*a vaca foi pro brejo*).

Até o momento vimos a arquitetura básica do modelo. Há muito a ser desvendado e desenvolvido a partir daqui. Esta seção, no entanto, não poderia ser encerrada sem que fosse deixado claro que todo o mecanismo derivacional da Morfologia Distribuída se articula em torno de 3 propriedades centrais. Apesar de elas já terem sido anteriormente apresentadas, ainda que de maneira informal, cabe retomá-las aqui como uma síntese do modelo.

Inserção Tardia: a sintaxe opera apenas com categorias abstratas, sem expoentes fonológicos. Somente após o fim da

derivação sintática é que a fonologia é inserida na estrutura, com base no inventário de Itens de Vocabulário disponíveis na Lista 2.

Subespecificação de Itens de Vocabulário: os Itens de Vocabulário não precisam ser especificados para todos os traços encontrados em um nó terminal sintático.

Estrutura Sintática por Toda a Derivação: a derivação sintática atua na formação tanto de palavras quanto de sentenças, sem haver – a princípio – uma distinção clara entre essas duas classes, a não ser após a inserção de material fonológico.

3. Remarks como origem do anti-lexicalismo

Após apresentarmos o modelo da Morfologia Distribuída e seu contraponto com a Hipótese Lexicalista, trazemos aqui uma interessante controvérsia.

Conforme foi dito anteriormente, Chomsky (1970) é comumente considerado como o berço da Hipótese Lexicalista. Porém, Marantz (1997), um dos proponentes da Morfologia Distribuída, apresenta uma visão um tanto diferente sobre a ideia que estava em jogo no texto de Chomsky (1970). Para o autor, a proposta de Chomsky (1970) não é uma defesa do léxico como componente formador de palavras, tal como se desenvolveu após a Hipótese Lexicalista.

Muito do que se discute em Chomsky (1970) está em torno de dados como a seguir:

- (13) a. *The tomatoes grow.*
 “Os tomates crescem.”

b. *John grows tomatoes.*

“João cultiva tomates.”

c. *The growth of tomatoes.*

“O crescimento dos tomates.”

d. **John's growth of tomatoes.*

“*O crescimento de João dos tomates”

O que esses dados revelam é uma assimetria entre a forma verbal e a forma nominal de *grow*. Enquanto o verbo *grow* pode tomar um único argumento (ex. (13)a), ou dois argumentos (ex. (13)b), a forma nominal, *growth*, pode tomar apenas um argumento (ex. (13)c), paralelamente a (13)a, sendo agramatical o paralelo nominalizado de (13)b, com dois argumentos, como visto na sentença (13)d.

Marantz enfatiza uma característica bastante importante da solução que Chomsky denomina *lexicalista*: nominalizações devem ser construídas a partir de entradas lexicais sem informação categorial.

Essa informação indica uma visão de que as entradas lexicais possuem características semânticas intrínsecas que dizem respeito, por exemplo, aos argumentos que podem tomar. No entanto, essas mesmas entradas lexicais não estão relacionadas por si só a categoria alguma (nome, verbo, etc.). Em outras palavras, quando a entrada *grow* está armazenada no léxico, há a informação sobre o número de argumentos que ela pode tomar, mas não há informação sobre se ela é um nome ou um verbo.

Uma entrada lexical como *grow* pode, então, combinar-se livremente em um ambiente verbal ou em um ambiente nominal. E, caso seja combinada em um ambiente nominal, não poderá sofrer derivações características de ambientes verbais. Caso seja combinada em um ambiente verbal, não poderá sofrer derivações características de ambientes nominais.

Um tipo de derivação característica de ambiente verbal é a causativização, que permite que um argumento causador seja adicionado ao evento denotado pelo verbo. Dessa forma, se uma entrada lexical pode tomar apenas um argumento, nada impede que ela seja causativizada em um ambiente verbal e a construção resultante ganhe outro argumento.

Considerando que *grow* é uma entrada lexical com um único argumento, podemos explicar as ocorrências em (13) com o seguinte raciocínio. Em um ambiente verbal, *grow* tem a possibilidade de ser causativizado. Como explicado acima, isso permite que um argumento causativo seja inserido na construção. Esse novo argumento não é entendido como argumento da entrada lexical *grow*, mas de um evento de causação adicionado no ambiente verbal. Isso possibilita a ocorrência da entrada lexical *grow* tanto em contextos verbais intransitivos (ex. (14)a) como transitivos (ex. (14)b):

- (14) a. *The tomatoes grow.*
 “Os tomates crescem.”
 b. *John grows tomatoes.*
 “João cultiva tomates.”

Em ambientes nominais, no entanto, como não há a possibilidade de causativização, *grow* pode apenas combinar-se com um único argumento.

- (15) a. *The growth of tomatoes.*
 “O crescimento dos tomates.”
 b. **John’s growth of tomatoes.*
 “*O crescimento de João dos tomates”

Em (15)a *grow* se combina com o argumento *tomatoes*, um

único argumento como o esperado. Em (15)b, no entanto, *grow* se combina com dois argumentos, *tomatoes* e *John*. Como o ambiente das sentenças em (15) é nominal, não havendo a possibilidade de causativização, o mesmo número de argumentos previsto pela entrada lexical (um único, no caso de *grow*) deve ser realizado na construção, impossibilitando a ocorrência de mais de um argumento.

O fato de que nominalizações com o sufixo *-ing*, gerundivas, podem ocorrer com dois argumentos pode ser tomado como evidência para o raciocínio apresentado aqui e para o caráter verbal desse tipo de nominalização:

(16) *John's growing tomatoes.*

“Os tomates que João está cultivando”.

As nominalizações em *-ing* são caracterizadas como nominalizações deverbais. Ou seja, para compor uma nominalização gerundiva, uma entrada lexical precisa ser combinada em ambiente verbal para depois ser nominalizada por *-ing*. Dessa forma, é possível que *grow* combinado em um ambiente verbal com causativização derive a sentença em (16) acima.

O que Marantz (1997) lê a partir de Chomsky (1970), portanto, é que a solução chamada de *lexicalista*, na verdade, só tem esse nome por supor (i) um componente com uma listagem de entradas lexicais com propriedades de seleção de argumentos e (ii) que as transformações devem obedecer a essas propriedades e não se aplicar de forma irrestrita para todas as entradas. Ou seja, o nome *lexicalista* tem mais relação com a adição de uma propriedade seletional nas entradas lexicais do que com a introdução de um componente lexical com suas próprias regras.

RESUMINDO

Neste capítulo abordamos o percurso da morfologia na Teoria Gerativa, com especial ênfase para o surgimento da Hipótese Lexicalista e para a forma como a Morfologia Distribuída se contrapõe à visão de que a Gramática apresenta dois componentes gerativos.

É importante ter em conta que, nos inícios da Teoria Gerativa, a morfologia era um nível de análise com pouca autonomia, estando os processos relacionados à formação de palavras associados ora a transformações sintáticas, ora a processos fonológicos. Chomsky (1970) é associado a esse movimento de separação da morfologia, sendo considerado o precursor da Hipótese Lexicalista, muito embora haja controvérsias.

A Hipótese Lexicalista foi desenvolvida ao longo dos anos 1970, com especial destaque para os trabalhos de Jackendoff (1975) e Wasow (1977). No entanto, o primeiro modelo completo de léxico voltado aos estudos da morfologia é visto em Aronoff (1976).

A Morfologia Distribuída, por outro lado, vê na flexibilização da correspondência entre os átomos de som e os átomos de significado uma oportunidade de se extinguir o léxico gerativo. São três as propriedades básicas da Morfologia Distribuída: inserção tardia de material fonológico, subespecificação de itens de vocabulário e estrutura sintática por toda a derivação.

Não faltarão oportunidades, na leitura daqui em diante, para compreender melhor cada propriedade do modelo. Mas é muito importante que o leitor tenha em conta que a Morfologia Distribuída não é uma teoria pronta, um modelo fechado. Muitas dessas propriedades são discutidas e partes do modelo vêm sendo aprimoradas desde o seu lançamento.

PARA SABER MAIS

Indicamos aqui algumas leituras para maior aprofundamento das questões que estão na base do modelo da Morfologia Distribuída. A primeira leitura, indispensável para tal aprofundamento e para a compreensão dos problemas de interface morfologia-sintaxe que estavam sendo investigados nos inícios da Teoria Gerativa, não poderia ser outra que não o texto *Remarks on nominalizations* de Noam Chomsky. Sobre a Hipótese Lexicalista, recomenda-se a leitura de *The lexicalist approach to word-formation and the notion of the Lexicon*, de Sergio Scalise e Emiliano Guevara. Trata-se de um texto que aborda o lexicalismo a partir da morfologia e explica os diferentes modelos de Morfologia ao longo da história da Teoria Gerativa. O pensamento de Morris Halle acerca da formação de palavras fundamenta muito da Morfologia Distribuída. Seu texto *Prolegomena to a theory of word formation* sintetiza muitas de suas concepções e ajuda a entender alguns porquês da Morfologia Distribuída. Alguns dos problemas mais intrincados de interface entre morfologia e sintaxe são abordados em *Where's morphology?* De Stephen Anderson. O texto também fundamenta uma característica central da Morfologia Distribuída, a *Inserção Tardia*. Por fim, o artigo *No escape from syntax: don't try morphological analysis in the privacy of your own lexicon*, de Alec Marantz, revisita *Remarks on nominalizations* sob um olhar não-lexicalista.

EXERCÍCIOS

Exercício 1. Indique com M as características associadas à Morfologia Distribuída e com L as características associadas à Hipótese Lexicalista:

a. Morfologia derivacional é responsabilidade do léxico.

- b. A derivação sintática manipula raízes sem especificação categorial.
- c. A derivação sintática lida com palavras e alguns morfemas flexionais.
- d. Na derivação sintática, um morfema tem um som e um significado associado a ele.
- e. As palavras são formadas ao longo de toda a derivação na Gramática, não havendo um componente específico para isso.

Exercício 2. Considere os dados do norueguês (bokmål):

- a) Masc: gutt – menino
Fem: jente – menina
Neut: hus – casa
- b) Masc: gutten – o menino
Fem: jenten – a menina
Neut: huset – a casa
- c) Masc: gutter – meninos
Fem: jenter – meninas
Neut: hus – casas
- d) Masc: guttene – os meninos
Fem: jentene – as meninas
Neut: husene – as casas

Tendo em conta o que foi apresentado sobre os Itens de Vocabulário, com apenas os traços [animado], [plural], [definido], depreenda até quatro regras que consigam dar conta da distribuição dos morfemas nos dados acima.

