

REVEL NA ESCOLA: MORFOLOGIA DISTRIBUÍDA

Ana Paula Scher¹

Universidade de São Paulo

Os processos de formação de palavras nas línguas naturais, de modo geral, sempre chamaram a atenção dos interessados em saber mais sobre como as palavras que nós conhecemos, em nossa língua materna e em outras, se tornam o que elas são. Alguns se interessam por seus aspectos etimológicos, diacrônicos, enquanto outros têm mais interesse pelo caráter sincrônico dessas formações, pelo que elas são em um dado momento do processo evolutivo de uma certa língua. Alguns se preocupam mais intensamente em desvendar as relações semânticas que se estabelecem entre palavras diferentes de uma mesma família, enquanto outros se voltam para questões concernentes à sua estrutura interna. Há, portanto, diferentes formas de se abordarem as questões relativas aos processos de formação de palavras nas línguas naturais.

O contraste que vamos marcar aqui, no entanto, tem origem, curiosamente, em um ponto em comum para os grupos contrastantes. Entre os estudiosos dos processos de formação de palavras, há os que procuram identificar as propriedades de uma arquitetura gramatical e, dentro dela, o lugar em que se formam as palavras. Nesse sentido, como há diferentes componentes nessa arquitetura gramatical, esse grupo pode ser dividido em pelos menos dois subgrupos de pesquisadores. Assim, há os que defendem um modelo *lexicalista* de formação de palavras, partindo do pressuposto de que as palavras se formam por meio de um conjunto de regras lexicais, independentes e diferentes das regras sintáticas da gramática, assumindo, explicitamente, que o processo de formação de palavras **não** se dá por meio de

¹ Doutora em Linguística pela Unicamp. Professora doutora da Universidade de São Paulo – USP. Bolsa Pq2 CNPq: 312610/2013-0.

transformações sintáticas. Para esses, então, a arquitetura da gramática deverá dispor de componentes distintos que se ocupem, por um lado, da formação de palavras, o *Léxico*, e, por outro, da formação de sentenças, a *Sintaxe*. No outro lado da moeda, há os que assumem exatamente o contrário, ou seja, os que pensam que morfologia é sintaxe, e que as mesmas regras que formam sentenças se ocuparão da formação de palavras. Essa forma de pensar tem a consequência imediata e desejável de reduzir a parafernália necessária para compor a arquitetura da gramática, já que um único componente, a *Sintaxe*, poderá dar conta da formação de palavras e sentenças. Esse é o modelo de que falaremos neste *ReVEL na Escola: a Morfologia Distribuída*.

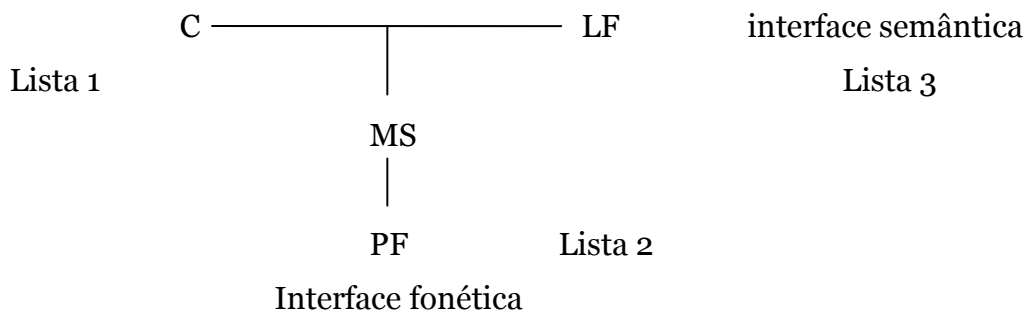
Trata-se de uma abordagem relativamente recente, formulada por Morris Halle e Alec Marantz (H&M, daqui para frente) e publicada em 1993 para a investigação dos processos de formação de palavras nas línguas naturais que procura demonstrar que a arquitetura da gramática pode ser mais simples do que se vem sugerindo por pesquisas que assumem a divisão de trabalhos entre *Léxico* e *Sintaxe*. Neste ensaio, procuraremos apresentar as propriedades do modelo da Morfologia Distribuída, MD, de agora em diante, realçando as que mais têm a ver com a ideia do *ReVEL na Escola*, trazendo exemplos simples de algumas línguas naturais para ilustrar o seu funcionamento.

O modelo da MD proposto em H&M (cf. também Marantz, 1997) se constitui em um modelo de análise morfológica baseado no morfema, que elimina o léxico, como um componente independente, da arquitetura da gramática e o substitui por listas distribuídas pelos demais componentes dessa gramática. Assim, existe uma lista 1, também chamada de *Léxico Reduzido*, uma lista 2, denominada *Vocabulário* e uma lista 3, conhecida como *Enciclopédia*.

Nesse modelo, não existe um componente lexical nos moldes do que se propõe em modelos lexicalistas, ou seja, um componente capaz de definir as propriedades de natureza formal, fonológica ou semântica dos itens lexicais e capaz, ainda, de operar sobre esses itens lexicais, com regras próprias, para formar unidades maiores. Essas mesmas propriedades, no entanto, estão presentes na arquitetura da gramática, distribuídas por seus outros componentes.

Em particular, o componente sintático (representado por C no diagrama em (1)) é o seu único componente gerativo², operando sobre o material de que é constituída **lista 1**.

(1) A organização da gramática³



Por ser a provedora do material que será manipulado pela sintaxe, a **lista 1** pode ser tomada como uma substituta do Léxico. Essa lista, determinada pela Gramática Universal e, talvez, por princípios específicos de cada língua, contém as raízes acategoriais da língua, além de feixes de traços gramaticais abstratos. É como se ela fosse um conjunto⁴ com os seguintes elementos, entre outros.

(2) Lista 1⁵: raízes e traços morfossintáticos abstratos

[raiz], [Det], [plural], [passado], [1pessoa], [CAUSE], etc⁶

A sintaxe manipula as raízes e os traços abstratos que vêm da lista 1. Os nós terminais de uma estrutura sintática são, portanto, feixes de traços formais, desprovidos de qualquer conteúdo fonológico ou idiomático. Esses feixes de traços constituem os ambientes variados em que podem ser inseridas as raízes e que vão ser os responsáveis pela categorização dessas raízes: a depender das camadas funcionais que as dominam, as raízes poderão ser categorizadas como verbos, nomes, adjetivos, etc.

² Nos termos dos idealizadores do modelo, *syntax all the way down*, que vimos traduzindo como *sintaxe morro abaixo*

³ C – Sistema Computacional; PF – Forma Fonética; MS – Estrutura morfológica; LF – Forma Lógica

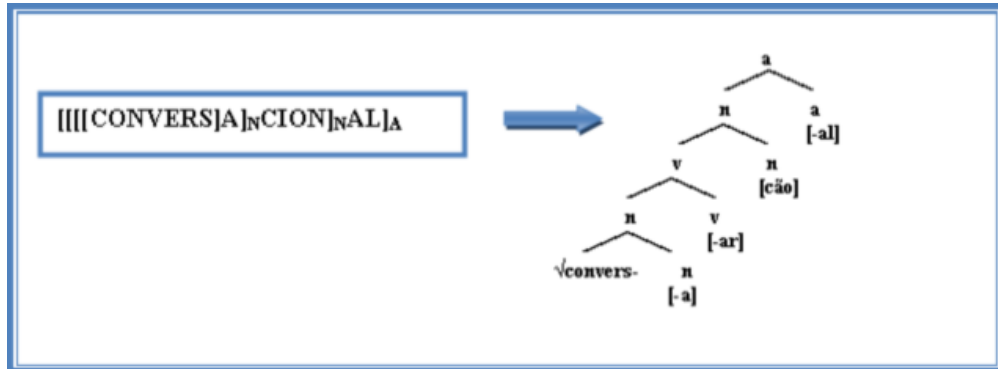
⁴ Deixemos de lado a discussão sobre se o termo *conjunto*, aqui, remete à Teoria de Conjuntos da matemática.

⁵ Esta é uma representação apenas parcial deste conjunto.

⁶ Esses exemplos de elementos pertencentes à lista 1 estão em Harley & Noyer (1999).

Um exemplo do que ocorre na sintaxe vem a seguir, com a derivação de uma palavra complexa do português, tal como *conversacional*:

(3) Derivação de *conversacional*⁷



O exemplo em (3) mostra o passo a passo da derivação sintática para a palavra *conversacional*, que começa com a concatenação da raiz acategorial $\sqrt{\text{CONVERS}}$ a um categorizador nominal (*n*), que deriva a estrutura sintática do nome *conversa*. O que vem a seguir são recategorizações sucessivas dessa raiz por meio de sucessivas concatenações que ocorrem nessa estrutura: insere-se, portanto, um categorizador verbal, na sequência, outro categorizador nominal e, finalmente, um categorizador adjetival, derivando, respectivamente, as estruturas sintáticas do verbo *conversar*, do nome *conversação* e, finalmente, do adjetivo *conversacional*.

Toda essa derivação sintática é abstrata, ou seja, desprovida de fonologia, já que os elementos com os quais a sintaxe trabalha, vindos da lista 1, não contêm material fonológico. Ao final dos trabalhos da sintaxe, a estrutura sintática se ramifica e envia parte desse conjunto de traços para a forma lógica (LF), para interpretação, a outra parte sendo enviada para a estrutura morfológica (MS) e, em seguida, para a forma fonética (PF).

Em algum lugar no caminho entre MS e PF, os traços abstratos da derivação sintática se realizam por meio da *inserção tardia*, nos nós terminais dessa derivação, de material fonológico, previsto por um conjunto denominado *Vocabulário*, que constitui a **lista 2**. Uma operação denominada *Spell-out* associa itens de vocabulário, ou seja, informações fonológicas, aos morfemas abstratos ou nós terminais que a sintaxe derivou. É só nesse momento que são inseridos, portanto, os itens de

⁷ Este exemplo de derivação aparece na página do GREMD – Grupo de Estudos em Morfologia Distribuída da USP <http://gremd.fflch.usp.br/node/8>.

vocabulário [-a], [-ar], [-ção] e [-al], que aparecem em (3) e que representam a realização fonética dos morfemas abstratos daquela estrutura.

Algum reajuste fonológico pode ser necessário depois dessa inserção tardia de material fonético e isso vai acontecer de acordo com as regras fonológicas da língua em questão. No caso em exame, esse tipo de regras se encarregará de externar a fonologia *-acional*, em lugar de *aaçãoal*, que representa a simples junção do material fonético correspondente aos morfemas categorizadores relevantes.

No caso mais simples, a relação entre os itens de vocabulário e os morfemas abstratos é de um para um: um único item de vocabulário poderá ser associado ao morfema abstrato correspondente⁸. No entanto, a aplicação de *Spell-out* a morfemas-*f* (do tipo funcional), por exemplo, pode gerar uma competição entre os itens de vocabulário compatíveis com o morfema abstrato em questão por um motivo muito simples. Enquanto os traços formais desse morfema, que foi derivado pela sintaxe, são completamente especificados, os do item de vocabulário, que indicam o tipo de morfema abstrato que eles podem preencher, são subespecificados em vários graus. A formação do plural do inglês pode ser um exemplo interessante aqui:

(4) singular	plural
ox	oxen
cow	cows

(5) Itens de Vocabulário

/-en/	↔	___ /ox, child, etc.../
/-s/	↔	___ / NDA

As especificações dos dois itens de vocabulário em (5) deixam claro que, para o nome *ox*, há uma marca específica que realizará o traço [pl], nomeadamente, *-en*, derivando a forma *oxen*. Por outro lado, na formação do plural do nome *cow*, o traço [pl], se realizará como a forma *default* *-s*, que se aplica a todos os ambientes para os quais não há uma opção mais específica, uma vez os itens de vocabulário não revelam nenhuma forma específica particular consistente com o traço [pl] no contexto do

⁸ Diversos fatores podem alterar essa relação, no entanto: fissão de morfemas, remoção de traços morfossintáticos por empobrecimento ou deslocamentos locais de itens de vocabulário por concatenação de morfemas, além de inserção pós-sintática de morfemas dissociados.

nome *cow*. Dessa maneira, o plural de *cow* se realizará como *cows*. O que se vê, portanto, é que a competição entre os itens de vocabulário possíveis para determinados traços gramaticais define e insere na estrutura, pós-sintaticamente, o mais completamente especificado, cujos traços sejam compatíveis com aqueles do morfema abstrato.

Além do *Léxico reduzido* (lista 1) e do *Vocabulário* (lista 2), uma *Enciclopédia* (**lista 3**) liga os *Itens de Vocabulário* a seus significados. Essa lista define significados especiais para determinadas raízes, dependendo do contexto sintático dessas raízes, dentro de um domínio local definido sintaticamente pela projeção de um núcleo sintático.

CONSIDERAÇÕES FINAIS

A breve exposição de algumas das propriedades do modelo da MD neste espaço denominado *ReVEL na Escola* não permite uma exploração detalhada das potencialidades do modelo. É possível deduzir, no entanto, do material apresentado aqui de forma tão concisa, que uma proposta como essa suscita inúmeras questões que precisam ser respondidas em face das novidades que o modelo representa. Será inevitável a rediscussão de temas tão caros ao domínio dos processos de formação de palavras, e já exaustivamente investigado em domínios lexicalistas, tais como os tradicionais processos derivacionais, flexionais, composicionais, além dos processos não-concatenativos de formação de palavras, dentro desse domínio não-lexicalista, representado, aqui, pela MD. O GREMD da USP (cf. nota 7) vem fazendo isso há cerca de 10 anos e tem feito descobertas interessantes, ao mesmo tempo que vem contribuindo para que o desenho desse modelo seja refinado e chegue ao que se considera um desenho ótimo.

REFERÊNCIAS

1. GUEVARA, E.; SCALISE, S. "The lexicalist approach to word-formation and the notion of the lexicon". In: STEKAUER, P.; LIEBER, R. *Handbook of Word-Formation*. Amsterdam: Springer, 2005, p. 147-187.

2. HALLE, M.; A. MARANTZ. 1993. "Distributed Morphology and the Pieces of Inflection". In: HALE, K.; & KEYSER, S. J. (eds.), *The View from Building 20: Essays in Linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press.
3. HARLEY, H.; R. NOYER. 1999. "State-of-the-Article: Distributed Morphology". *Glot* 4.4: 3-9.
4. MARANTZ, A. 1997. "No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon", *University of Pennsylvania Working Papers in Linguistics* 4: 201-225.