



**UNIVERSIDADE FEDERAL DA BAHIA**  
**INSTITUTO DE LETRAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS E LINGUÍSTICA**

**BRAULINO PEREIRA DE SANTANA**

**INTERSECÇÕES ENTRE MORFOLOGIA E LÉXICO**  
**NO CONCEITO DE ENTRADA LEXICAL**

**Salvador-Bahia**

**2010**

**BRAULINO PEREIRA DE SANTANA**

**INTERSECÇÕES ENTRE MORFOLOGIA E LÉXICO**

**NO CONCEITO DE ENTRADA LEXICAL**

Tese apresentada ao Programa de Pós-Graduação em Letras e Linguística da Universidade Federal da Bahia como requisito parcial para a obtenção do grau de Doutor em Linguística.

**Orientadora: Dra. Serafina Maria de Souza Pondé**

**Salvador-Bahia**

**2010**

**Sistemas de Bibliotecas - UFBA**

Santana, Braulino Pereira de Santana.

Intersecções entre Morfologia e Léxico no Conceito de Entrada Lexical / Braulino Pereira de Santana  
247 f.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Serafina Maria de Souza Pondé.

Tese (doutorado) – Universidade Federal da Bahia, Instituto de Letras, Salvador, 2010.

1. Entrada lexical. 2. Morfologia. 3. Léxico. 4. Semântica lexical. 5. Unidades lexicais. 6. Lexemas. 7. Projeção semântica. 8. Formação de palavras. 9. Agente. I. Pondé, Serafina Maria de Souza. II. Universidade Federal da Bahia. Instituto de Letras. III. Título.

CDD – 817

CDU – 87.402

A

Iolanda Pereira de Santana

A

Priscila Maynard Araújo, Suani Vasconcelos, Vilbégina Monteiro, Lídia Nunes Cunha, José Amarante Sobrinho, José Imperatriz dos Santos, João Batista de Castro Júnior, João F. Barreto, Ricardo Nascimento Abreu, Marcos Aurélio dos Santos Souza, Freddy Isbister, Francis Augusto Medeiros, Wilson Gabriel de Jesus, José Bernardo e José Rodrigo.

## **AGRADECIMENTOS**

Serafina Pondé é uma pessoa que se confunde com as melhores coisas que já me aconteceram na vida: este trabalho só foi possível com a sua inestimável inteligência.

A aventura linguística humana na Terra começa com uma simples Entrada Lexical.

## RESUMO

Uma entrada lexical, unidade semântica que engloba afixos, palavras e metáforas, é o primeiro contato que uma criança em processo de aquisição tem com a sua língua materna. Este trabalho centra seu foco no conceito de entrada lexical e suas intersecções com a Morfologia e o Léxico. Rastreia uma série de pressupostos teóricos tendo como foco a problematização de três questões: quais intersecções possíveis uma entrada mantém com o léxico e com a morfologia; como a mente humana absorve e estoca entradas numa língua; e qual a natureza semântica do conhecimento dos falantes que os faz ter consciência e intuições sobre o que significa uma entrada lexical. Questões tão amplas como essas exigem de um pesquisador um projeto investigativo que impõe não um conjunto criterioso de dados, mas um entrelaçamento de um arcabouço teórico que conjugue língua e mente. Não investimos em processos de uso neste trabalho, mas focamos processos cognitivos que respondem por checagens na mente dos falantes a respeito de como eles preparam as entradas com as quais se comunicam, como eles conseguem entradas a partir de elementos idiossincráticos e padrões de organização, e, por fim, quais estratégias os falantes utilizam para estocar durante toda a sua vida as entradas em suas mentes. A aventura linguística humana na Terra dispara com uma simples Entrada Lexical.

Palavras-chave: Entrada Lexical. Morfologia. Léxico. Checagem. Projeção semântica.

## **ABSTRACT**

A lexical entry is a semantic unit which comprehends affixes, words and metaphors. It is the first contact that a child, during its language acquisition process, has with its mother tongue. This work emphasizes the lexical entry concept and its intersection with Morphology as well as the lexical meaning of words. It also uses several theoretical premises which focuses on three problems: 1. Possible common points that a lexical entry has with the lexical meaning of words and Morphology; 2. how the human mind absorbs and keeps lexical entries in any language and 3. The semantic nature of the speakers' knowledge that makes them aware of a lexical entry meaning. Such problems demand an investigative process that relates theories which take into consideration not only the language itself but also human mind. In this work we haven't taken into consideration everyday language but we have focused in the cognitive processes which respond to checking within the speakers' mind concerning the way they prepare these lexical entries to take part in the communication process. Moreover, it is also concerned not only about the way speakers get these lexical entries from the idiosyncratical elements and patterns of organization but also the strategies used by the speakers to have these lexical entries throughout all their lives in their minds. The human linguistic adventure on earth begins with a simple lexical entry.

Key-words: Lexical entry. Morphology. Lexical meaning. Checking. Semantic projection.



## LISTA DE QUADROS

Quadro 1: Caracterização de unidade lexical

Quadro 2: Usos de uma entrada lexical

Quadro 3: Modulação de traços semânticos

Quadro 4: Normalização

Quadro 5: Processo de aquisição de linguagem

Quadro 6: Adquirindo novas entradas lexicais

Quadro 7: Itens de conteúdo e itens relacionais

Quadro 8: Postulado de regras

Quadro 9: Unidade conceitual e totalidade interna

Quadro 10: Holismo e localismo

Quadro 11: Teste de ambiguidade

Quadro 12: Relações paradigmáticas

Quadro 13: Descrição de traços gramaticais

Quadro 14: Operações de merger

Quadro 15: Checagem semântica

Quadro 16: Extensão e intensão

## SUMÁRIO

<b>1 PRÓLOGO - <i>O limbo lexical</i></b>	<b>10</b>
<b>2 INTRODUÇÃO</b>	<b>13</b>
<b>3 A NOÇÃO DE ENTRADA LEXICAL NA MORFOLOGIA E NO LÉXICO</b>	<b>17</b>
3.1 QUANDO A MORFOLOGIA SE APROPRIA DO LÉXICO	29
3.2 ENTRADA LEXICAL E PRODUTIVIDADE MORFOLÓGICA	31
<b>4 ENTRADA LEXICAL E DISTINÇÃO ENTRE UNIDADES LEXICAIS E LEXEMAS</b>	<b>35</b>
4.1 CONSTITUINTES SEMÂNTICOS	43
4.2 CRIATIVIDADE GRAMATICAL E LÉXICO	51
4.3 A SIGNIFICATIVIDADE INFORMACIONAL E COGNITIVA	61
4.4 DELIMITANDO LEXEMAS	66
4.5 O PRINCÍPIO DAS RELAÇÕES RECORRENTES	84
<b>5 SELEÇÃO E MODULAÇÃO DE SENTIDO</b>	<b>101</b>
5.1 CONSEGUINDO ENTRADAS LEXICAIS	113
5.2 FORMAÇÕES RETROATIVAS	126
<b>6 A NOÇÃO DE ENTRADA LEXICAL NA ESTRUTURA ARGUMENTAL</b>	<b>131</b>
6.1 ITENS DE CONTEÚDO E ITENS RELACIONAIS	134
6.1.1 Estrutura argumental	143
6.2 REGRAS DE PROJEÇÃO SEMÂNTICAS E ENTRADAS LEXICAIS	154
<b>7 ENTRADA LEXICAL E DERIVAÇÃO</b>	<b>164</b>
7.1 DERIVAÇÃO E LÉXICO	165
7.1.1 Entradas-sufixos como unidades lexicais e lexemas	177
7.1.2 Definindo um sufixo-lexema e um sufixo-unidade lexical	178
7.1.3 A ambiguidade das entradas-sufixos	191

<b>8 A CHECAGEM DAS ENTRADAS LEXICAIS</b>	<b>205</b>
8.1 A LÓGICA DA FLEXÃO	212
8.2 CRITÉRIOS DE DISTINÇÃO ENTRE DUAS MORFOLOGIAS: A MORFOLOGIA DA FORMAÇÃO DE PALAVRAS E A MORFOLOGIA DOS PARADIGMAS FLEXIONAIS	216
<b>9 CONCLUSÃO</b>	<b>229</b>
<b>10 REFERÊNCIAS</b>	<b>234</b>

## 1 PRÓLOGO - *O limbo lexical*

O léxico é uma lista de palavras que jaz na mente dos falantes. Quando afirmamos isso, estamos deduzindo que, no cérebro dos seres humanos, há um dicionário, um inventário de palavras. Dicionários como o Houaiss e o Aurélio, que inventariam um rol de palavras da língua portuguesa, podem ser tomados como uma lista pública, e a lista que pressupomos haver na mente dos falantes é parecida com a lista inventariada por esses dicionários.

Em nossas mentes, há palavras guardadas das quais pouco nos lembramos; há palavras que são usadas com mais frequência do que outras; há palavras que mantemos sob censura; há palavras que guardamos há muito tempo; não existem algumas palavras, que podem fazer parte da mente de um falante e não constar da mente de outros, assim como há palavras no Aurélio que não estão no Houaiss e vice-versa.

Conhecer uma lista de palavras não é, no entanto, suficiente para por a língua que dominamos em uso: é apenas o ponto de partida. Somos dotados de capacidade genética para jogar com as palavras na construção das frases que acionamos no dia a dia da comunicação por intermédio de uma língua, ou seja, dominamos regras de combinação de palavras.

O jogo começa quando retiramos uma palavra do nosso dicionário mental e a dispomos numa frase. Suponhamos que estamos procurando uma palavra no dicionário Aurélio para a utilizarmos numa frase de um texto que estamos escrevendo. Essa palavra é um verbo. Como sabemos, os verbos estão dispostos no Aurélio em sua forma infinitiva, tal como o verbo [amar]. Para utilizarmos esse verbo na frase que estamos elaborando, é necessário que ele sofra um “ataque gramatical”. Numa frase como *Joana ama Márcia*, o verbo *amar* sofreu um ataque gramatical nos seguintes termos: foi conjugado no presente; assumiu a terceira pessoa para concordar com o sujeito *Joana*; está no singular, pois o sujeito também está no singular; assumiu uma forma declarativa, ao invés de uma forma condicional, como na frase *Que Joana ame...* Temos como hipótese que, na mente, quando acionamos uma palavra para uso, ocorra um processo parecido: as palavras estão em estado de dicionário – no nosso caso aqui, estado de dicionário mental – sacamos a palavra, empreendemos ataques gramaticais, como exposto acima, e construímos frases.

Vamos pressupor que na mente exista um *limbo lexical*. Um lugar onde as palavras estão em repouso – parecido com o repouso das palavras no Houaiss e no Aurélio – para que sejam acionadas quando precisarmos delas em algum momento. Adormecidas em nosso cérebro, são acordadas e dispostas para o uso a partir de determinados mecanismos mentais. Vamos propor dois deles aqui: o princípio da projeção semântica e o princípio da projeção gramatical.

As palavras são projetadas, alçadas, retiradas do limbo lexical e dispostas numa frase. Uma palavra como *banco*, por exemplo, guardada no dicionário mental, é projetada numa sentença da língua portuguesa. No caminho que vai do limbo lexical em direção a uma frase, é preciso que ela chegue “inteira” semanticamente. Se procurarmos no dicionário Aurélio o significado básico dessa palavra, não seremos capazes de identificá-lo, dado que, para ela, o Aurélio elenca uma série de sentidos: instituição financeira; um móvel sem braço onde as pessoas sentam; um verbo, como na frase *eu banco essa mulher*; um empecilho para navegantes, como na frase *um banco de areia encalhou o barco*. Vamos chamar esses vários sentidos de traços semânticos.

Ora, para que ela seja projetada numa frase como *Serafina sentou-se no banco da praça e deu comida aos pombos*, somente um desses traços é projetado, qual seja, o traço *móvel sem braço onde as pessoas sentam*. O falante, portanto, aciona esse traço e apaga todos os outros para delimitar aquele que seja necessário para construir a frase com a personagem Serafina. Os outros ficam em repouso no limbo lexical, em forma latente, e serão acionados em uma outra frase, para um outro contexto de uso.

Observamos até agora o princípio de projeção semântica sendo acionado. O princípio de projeção gramatical será acionado de forma parecida. A palavra *banco* será usada como substantivo na sentença com a personagem Serafina, e não como um verbo, como na sentença acima *eu banco essa mulher*; será acionado no singular, diferente de sentenças como *Serafina viu os bancos da praça, e eles são feitos de alumínio*. A gramática vai acionar traços que os linguistas chamam de categoriais: número; classe gramatical: substantivo, verbo etc; enquanto a semântica vai acionar traços relacionados ao significado da palavra.

Os mecanismos de projeção gramatical atuam acionando padrões, regularidades, como por exemplo: há uma regularidade em termos de número em língua portuguesa – a ausência ou presença de um [s] é capaz de dizer se uma palavra está no singular ou no plural. Como também de irregularidades: o gênero em língua portuguesa é altamente irregular, dado que o falante terá que aprender se o gênero de uma palavra é masculino ou feminino palavra a palavra, já que somente 4,5% das palavras da língua portuguesa mantêm regularidade como no par *menino/menina*. O restante dos substantivos é de gênero único, como *a tribo, o pente, a máquina, a parede* etc.

Ao projetar palavras numa frase é certo que os falantes têm conhecimento de duas coisas: eles sabem que possuem uma lista de palavras em suas mentes, e sabem também que

dominam regras de combinação dessas palavras nas frases. Os falantes, todos e cada um deles, são iguais em termos de trabalho de projeção semântica e gramatical de palavras na confecção de frases.

## **2 INTRODUÇÃO**

Dando vazão a intuições antigas, Rocha (1998, p.112) define um item ou uma entrada lexical como “uma forma linguística que o falante conhece ou utiliza” e assegura que as entradas lexicais constituem o léxico de uma língua.

A natureza simples dessa definição esconde por trás questões delicadas que vêm sendo enfrentadas ao longo da história da linguística contemporânea envolvendo léxico, morfologia e palavra. Este trabalho propõe enfrentar basicamente três dessas questões:

- a) Quais intersecções possíveis uma entrada mantém com o léxico na globalidade, e com a morfologia?
- b) Como o léxico absorve e estoca entradas numa língua?
- c) Qual a natureza do conhecimento semântico dos falantes que os faz ter consciência, e por vezes intuições, sobre o que seja uma entrada lexical?

A literatura linguística talvez já apresente consensos amplos e definições precisas, cercada por um grande número de dados de várias línguas, a respeito de uma entrada lexical. Numa outra ponta, algumas posturas teóricas em linguística propõem, inclusive, abandonar esse conceito (por exemplo, PERINI, 1997; HALLE & MARANTZ, 1992), dada a sua suposta inutilidade teórica ou dificuldades em delimitar as suas fronteiras semânticas.

O modelo Morfologia Distribuída (MD), por exemplo, sustenta não haver Léxico, ao menos no sentido familiar como proposto pela Gramática Gerativa, em trabalhos dos anos de 1970 e 1980, ao rejeitar a Hipótese Lexicalista. Marantz (1997), por exemplo, argumenta, de forma bem sistematizada, uma proposta teórica contrária à noção de léxico gerativo. Nesses termos, noções como entrada lexical e léxico não têm significado para a teoria da MD, assim como lexicalização e entrada lexicalizada. Para ele, nada acontece no léxico sem que haja uma manipulação de traços gramaticais acionados pela sintaxe. Há uma rejeição do modelo lexicalista pois, se uma entrada é lexicalizada, isso corresponde a dizer que ela está naturalmente dada, e a sintaxe não teria sobre ela motivação alguma, como se ela fosse independente em interface com outros módulos da gramática, como a fonologia e a própria sintaxe.

Entretanto, vamos admitir a existência de entradas lexicais independentes da sintaxe. E vamos admitir que os falantes tenham intuições sobre estocagem dessas entradas, motivados por uma série de pressupostos que serão delineados a partir da abordagem em semântica lexical. No decorrer deste trabalho, vamos deixar bem claro como esses pressupostos podem ser acionados e como eles funcionam para a delimitação de entrada lexical. Estamos bem certos de que propomos uma visão regressiva sobre léxico, já que os falantes têm conhecimento o suficiente de sua língua para saber que, sem as entradas lexicais, não existiriam as frases.

Talvez gastar tempo, esforços mentais e recursos em torno de algo também supostamente supérfluo para a ciência linguística seja algo irrelevante ou até mesmo sem futuro. Apesar dessas posturas céticas, contudo, enfrentaremos as questões acima ancorados em alguns pressupostos teóricos fronteiriços entre morfologia e léxico.

Ao longo deste trabalho, investiremos em uma delimitação nuclear sobre léxico: *o léxico mental e o léxico real*. E quem pensa que o debate em torno das relações entre léxico mental e real, permeado pelo conceito de entrada lexical, já esteja superado precisa pensar melhor, pois uma entrada lexical é um item que um falante conhece (léxico mental) e eventualmente usa (léxico real); e quando se trata de explicar o pensamento e o comportamento humanos, a possibilidade de começar esse debate por meio da concepção elementar do papel das palavras na vida humana ganha relevância, e chega a ser fundamental. Admitir a natureza das palavras como ponto de partida da aventura humana na Terra fascina, ao mesmo tempo em que impõe necessidades teóricas.

Vamos investir na delimitação de como uma entrada é concebida no léxico e na morfologia. Faremos uma espécie de limpeza de terreno semântica, a partir de duas perspectivas: quando a entrada é vista como lexema, e quando a entrada é vista como unidade lexical. Vamos investir também nas perspectivas teóricas da semântica lexical sobre o conceito de entrada. Ao longo das propostas teóricas, a argumentação será ilustrada por dados, levantados a partir de duas fontes:

**Fonte 1:** entradas registradas no dicionário Novo Dicionário Aurélio da Língua Portuguesa, 3ª Edição, Revista e Atualizada, 2004 (doravante, Aurélio). Entradas com os sufixos [-dor]; [-eiro]; e [-nte].

**Fonte 2:** o romance *O Coronel e o Lobisomem* (CL), do carioca José Cândido de Carvalho. Outros dados complementares foram cunhados a partir da nossa experiência como pesquisador, oriunda das nossas intuições como falante.

Não é demais ressaltar que paira ao longo deste livro uma reflexão discreta sobre aquilo que nos separa como seres humanos do restante da natureza: o fato de dominarmos, usarmos e criarmos *palavras* (além de termos intuições precisas sobre os padrões que montam e desmontam as cartilagens de sua delicada anatomia) para a fascinante experiência de comunicação singular dos seres humanos na face da Terra: a capacidade de falar e de entender uma língua.



### **3 A NOÇÃO DE ENTRADA LEXICAL NA MORFOLOGIA E NO LÉXICO**

FOCO: Baseados em concepções que circulam de forma ampla, vamos apontar neste capítulo intersecções e rearranjos entre morfologia e léxico sobre o que significa uma entrada lexical, e suas potencialidades conceituais.

Em variados estudos relativamente consensuais (BLOOMFIELD, 1933; CÂMARA JR., 1970; ARONOFF, 1976; BASÍLIO, 1980; VILLALVA 1998), a Morfologia, por ser parte da gramática, e lidar com questões estruturais, trata da estrutura interna de palavras simples e complexas (BLOOMFIELD, 1933; CÂMARA JR., 1970), reais e potenciais (ARONOFF, 1976; BASÍLIO, 1980; VILLALVA, 1998) de uma língua. Essas palavras até podem não existir ainda, mas todas precisam se adequar à estrutura morfológica da língua sob o risco de não conseguir existir como palavra.

O conceito mais comum sobre léxico diz respeito a uma lista de itens que existem na língua, itens que um falante precisa conhecer, e tem que estocar, por serem signos arbitrários, idiossincráticos e muitas vezes imprevisíveis: não presumíveis de alguma forma.

Muitos dos itens da lista são palavras, embora o léxico também contenha unidades maiores como expressões idiomáticas, lexemas e unidades lexicais, como talvez também unidades menores como afixos – vamos assumir provisoriamente que palavras, expressões idiomáticas, lexemas, unidades lexicais e afixos sejam tipos de entrada lexical.

Nesses termos, à primeira vista, a Morfologia pode parecer dar conta das regularidades, enquanto o léxico – das irregularidades. E como também pode-se imaginar os dois tendo muito pouco a ver um com o outro, já que a morfologia trata de palavras potenciais e o léxico somente de palavras que existem, itens concretos, aqueles listados num dicionário – real ou virtual. Na verdade, os dois sistemas têm muito a ver um com o outro, por três razões:

- a) A primeira é que eles servem ao mesmo papel numa língua: ambos lidam com entradas lexicais.
- b) A segunda razão é que a morfologia e o léxico são interdependentes. A morfologia, que forma palavras a partir de palavras, encontra as palavras com que opera (e suas bases) no léxico; assim como o léxico se beneficia dos princípios que o estruturam, princípios delineados em teorias morfológicas.
- c) A terceira é que o léxico e a morfologia se articulam em torno de elementos mínimos que compõem as entradas.

Muitas vezes, essas intersecções têm levado alguns linguistas a hierarquizar o léxico num nível mais acima, enquanto a morfologia num nível mais abaixo, e pressupor que a morfologia está no léxico (JENSEN e STONG-JENSEN, 1984; ARONOFF, 2001), embora, ao dizer isso, esses linguistas estejam usando o termo *léxico* num sentido mais amplo e diferente, significando a fonte de todas as entradas, reais e potenciais, ao invés de um sentido

estrito – uma lista de itens<sup>1</sup> não presumíveis que herdamos da gramática tradicional e de procedimentos das teorias estruturalistas *a la* Bloomfield.

Por se fixar num mesmo escopo de interesses, mas por abordá-lo a partir de metodologias próprias, podemos dizer que, mesmo tendo as entradas lexicais como objetos comuns de suas pesquisas e análises, morfologia e léxico recriam este mesmo objeto sob perspectivas diversas, até mesmo contrastantes.

Como acontece com quaisquer entidades que compartilham uma tarefa, a morfologia e o léxico nem sempre agem de forma a chegar a denominadores comuns em torno do objeto de suas pesquisas e análises: morfologia e léxico muitas vezes compartilham tarefas diferentes. Esse compartilhamento desempenha papel central no sistema mais amplo da língua.

Para captar a recriação de seu objeto, vamos considerar a atuação de um simples falante/ouvinte. Quando falamos de léxico sob a perspectiva do falante/ouvinte, estamos falando do léxico mental ou virtual, a lista de entradas irregulares de que o falante/ouvinte tem conhecimento, armazenada na memória, e precisa estocar de uma forma ou de outra.

Primeiramente, vamos investir na definição de itens que existem e itens potenciais em termos do léxico mental ou virtual. Diremos que qualquer entrada lexical faça parte do léxico mental de nosso falante/ouvinte, ou a lista de entradas irregulares componha o léxico mental, ou de fato exista em termos funcionais, e que nenhuma outra entrada mais exista para ele. Assim, um item que possui todos os critérios para ser uma entrada real numa língua, mas, por não estar armazenado no léxico mental do indivíduo, não existe para essa pessoa, embora possa existir para um outro falante/ouvinte ou para uma outra comunidade de fala.

A assertiva de Basílio (1980) e de Rocha (1998), de que um dicionário não é um bom lugar para se descobrir se uma palavra existe ou não, dá conta provisoriamente da existência ou não de um item: há num dicionário itens desconhecidos pelos falantes, e há itens conhecidos pelos falantes que não estão listados num dicionário, por exemplo.

O item não listado, tanto na memória do falante quanto num dicionário público como o *Dicionário Houaiss da Língua Portuguesa* (2001) (doravante, Houaiss), é uma entrada em potencial, e diremos que entradas potenciais complexas ou simples, bem formadas morfologicamente, são fornecidas pela morfologia, mas não pelo léxico. Nesses termos, a idéia convencional de que palavras que existem numa língua como o português, por exemplo, compreendem todas as palavras do dicionário Aurélio ou Houaiss, ou algum outro dicionário

---

<sup>1</sup> Item tem a mesma acepção que entrada lexical.

da língua portuguesa, não se aplica a esse modelo de léxico e morfologia. “A diferença entre palavras que existem e palavras em potencial é definida somente em termos do léxico do indivíduo e em termos da morfologia”. (ARONOFF, 2001, p. 238)<sup>2</sup>.

Ressalta-se, nessas alturas, que, mesmo que o falante/ouvinte tenha entrado em contato, falado ou ouvido (ou lido) uma entrada antes, se essa entrada não tiver sido estocada, contudo, no léxico mental da pessoa por alguma razão, então ainda se trata de algo em potencial, em vez de uma entrada que exista, até que o léxico mental a tenha capturado.

De início, uma entrada será estocada quando contiver ao menos um morfema com uma informação semântica compartilhada. Tome-se, por exemplo, a entrada ?[áua]. Ela não possui estrutura morfológica aparente e não existe no mundo de fala adulto – é uma entrada do mundo da fala infantil. Alguém que tenha ouvido essa (?) palavra, mesmo em um contexto em que o seu sentido esteja claro, deve armazená-la em sua memória a fim de usá-la novamente, só assim será estocada no léxico mental do ouvinte.

De uma forma similar, uma entrada complexa morfológicamente terá dificuldade em ser estocada se um pedaço dela for desconhecido para o ouvinte. Um exemplo desse tipo é [aguardente], composta por [água] mais [ardente]. Suponha-se que um dos seus componentes [ardente] seja desconhecido, mas o outro [água] não seja, assim, então, novamente, devemos memorizar a palavra em sua inteireza se quisermos reusá-la com o mesmo sentido, mesmo se pudermos deduzir seu sentido do contexto em que a ouvimos antes. Como raramente temos acesso aos mesmos contextos de funcionamento dos itens a cada vez que falamos ou ouvimos, essa suposta palavra de pedaço desconhecido deve ser colocada no léxico.

Todos os componentes estruturais de uma entrada devem ser conhecidos para que o seu sentido possa ser deduzido a partir deles, e possa, dessa forma, ser estocada. Estamos deixando de lado, por uma necessidade metodológica, concepções funcionais sobre a língua. Uma palavra será armazenada no léxico se o falante ouvinte cobrir uma rede de traços semânticos exaustiva para que ela seja útil mais adiante para ele, só assim a entrada fará parte do estoque e será armazenada no léxico.

Somente a organização e a estrutura não dão conta das necessidades cognitivas de armazenagem. Tome, por exemplo, uma entrada como [pé-de-moleque]. Tanto [pé] quanto [de] e [moleque] são familiares, mas o sentido da entrada completa não tem nada a ver com o sentido de suas partes, até mesmo aqui o nosso falante/ouvinte ideal precisa lançar mão do

---

<sup>2</sup> The difference between which words exist and which are potential is defined solely in terms of the individual's lexicon and morphology. Todas as citações em língua estrangeira, cujos textos originais estão em notas de rodapé, foram traduzidas por nós.

estoque lexical a fim de ter uma esperança de reusar a palavra: se uma palavra não é previsível a partir de seus pedaços constituintes, deve ser estocada no léxico.

Por contraste, considere a entrada [nacionalização], de circulação quase universal em variadas comunidades de fala pelo Brasil. A sua leitura semântica pode ser deduzida a partir de sua estrutura, a partir dos pedaços menores que a compõem. A soma semântica desses pedaços é capaz de se agrupar para formar um todo semântico. Ela pode ser desmontada em partes menores, mas que englobam um mesmo espectro semântico: se retirarmos [-ização], por exemplo, o restante [nacional] ainda assim preserva um traço semântico contido em [nacionalização], a entrada completa. Para capturar o significado de uma entrada como [comercialização], o falante empreenderá esforços morfológicos parecidos com aqueles que ele utiliza para a entrada [nacionalização], por exemplo. Nesses termos, a morfologia é acionada.

Podemos deduzir seus pedaços componentes, eles são, digamos, legíveis: [nação], e [-al] (que forma adjetivos a partir de substantivos); [-izar] (que forma verbos a partir de adjetivos em língua portuguesa); e logo depois [-ção] (que forma substantivos abstratos). O seu significado é, nas palavras de Aronoff (2001, p. 239), transparente, e pode ser parafraseado grosseiramente como *ato ou fato de fazer algo ser nacional, tornar-se ou tornar algo nacional*. A transparência admitida por Aronoff não advém simplesmente da somatória dos seus pedaços componentes, apesar de podermos dizer isso também, mas de uma conjunção hierarquizada de seus constituintes. Filtra-se o seu significado a partir de um trabalho de “fricção” semântica entre esses pedaços que se juntam para formar um todo.

Vamos a um outro dado para ilustrar como morfologia e léxico se complementam no estudo dos itens. O plural em línguas como o inglês e o português, por exemplo, marcado por meio de afixos indicativos de número – sua ausência em [jornalista], marcando o singular; sua presença em [jornalista-s], marcando o plural. O falante não precisa estocar o item em si (como ele faz quando estoca os itens de uma lista do léxico), mas aplicar um padrão (presença de afixo de plural X ausência de afixo de plural) para os itens. A aplicação do padrão é de interesse central em morfologia.

Sendo o sentido real de uma palavra não divergente de seu sentido previsível, deduzido a partir de seus pedaços componentes e de sua morfologia, não há necessidade de essa palavra ser listada no léxico como se ela fosse estranha ao falante, idiossincrática. Lidando com suposições, vamos dizer que o falante nunca tenha ouvido/lido antes a palavra

[nacionalização], mas já tenha estocado no léxico a palavra [nacional], a partir disso, ele pode assumir um significado para a palavra [nacionalização], já que o componente morfológico da gramática do falante/ouvinte está apto a processá-la inteiramente. A partir disso, essa entrada está pronta para entrar no léxico. Vemos que há uma passagem da morfologia ao léxico, nessa linha de argumentação. Em tempos relativamente recentes, estivemos às voltas com a entrada [malufar]. Ninguém precisou recorrer ao dicionário para presumir o seu significado – nem no dicionário encontraríamos essa entrada. A morfologia, portanto, o jogo com o padrão, foi acionada para estabelecer traços semânticos contidos nessa entrada.

A palavra será, portanto, falada, ouvida, e mais provavelmente descartada em todas as partes, talvez seja criada e descartada novamente, mas não necessariamente estocada, ao menos se usada em algum sentido especial que não seja previsível a partir da morfologia.

A partir de deduções como essas, foi possível empreender esforços nos anos setenta para estabelecer uma morfologia gerativa baseada em regras ou padrões, prever a possível criação de novas palavras, deduzir possíveis bloqueios de criação de outras e assim por diante.

Vimos o que cabe e o que não cabe à morfologia e ao léxico. Constatamos que a morfologia lida com regularidades e com potencialidades de criação de novas palavras a partir de padrões, como também lida com a inteligibilidade de entradas novas ou recentes ao se jogar com os padrões de uma língua, enquanto cabe ao léxico, como uma lista de entradas, lidar com as irregularidades e com os elementos idiossincráticos. Agora, como morfologia e léxico estão interrelacionados?

Para ver como eles se relacionam, vamos observar determinados casos em que tanto uma quanto o outro interagem para explicitar as potencialidades gramaticais de uma entrada lexical. Vamos começar com um caso simples, aquele da categoria gênero linguístico em substantivos do português.

A categoria gênero linguístico provém, na maioria dos casos, do léxico, em poucos outros casos, da morfologia. O gênero virá do léxico se não houver marca afixada ao substantivo que identifique o seu gênero, no caso, [-a] para o feminino, e [-o] para o masculino (há outras marcas de gênero em língua portuguesa, mas vamos etiquetá-las todas como traços disponíveis para essas marcas aqui). Nesse caso, vamos chamar de irregular, e será estocado lá no léxico mental por conta de sua irregularidade, como [o pente] e [a tribo], e virá da morfologia no caso de ser, digamos, regular, como [gat-a] e [gat-o].

Mas agora surge uma questão. Se uma palavra tem um gênero irregular estocado no léxico, por que ela também não tem um gênero regular, provindo da morfologia? Nesse último caso, como um falante sabe que não deve dizer \*[o tribo] ao invés de dizer [a tribo], ou

dizer ambas as formas? Ou por que o falante às vezes não diz uma forma, e outras vezes, a outra forma? Algo deve estar evitando a morfologia de produzir, por precaução, o gênero regular para a mesma palavra que existe no léxico, mas de gênero irregular.

O mesmo pode ser dito a respeito dos verbos irregulares do português. Uma pessoa que sabe que o passado de [fazer] é [fiz] (um fato que deve ser estocado no léxico) não dirá \*[fazi], embora uma criança ou alguém nos primeiros estágios de aquisição do português como segunda língua possa dizer \*[fazi] porque a criança ou o estudante não tenha aprendido ainda a forma [fiz]. Léxico e morfologia juntam suas forças para garantir que uma forma seja usada, e a outra nem faça parte da gramática do falante.

O falante, de alguma forma, filtra os itens em seu cérebro antes de pronunciá-los. Para a questão levantada por Aronoff (2001, p.239), “o falante/ouvinte de alguma maneira checa o léxico para ver se uma palavra está lá, e somente recorre à morfologia se não encontra nenhuma”<sup>3</sup> – vamos admitir que sim. Saber como isso acontece vai ser nossa preocupação nos parágrafos abaixo. Para isso, vamos nos basear no artigo Morfologia e Léxico, de Aronoff & Anshen (2001).

Podemos investir nessas questões rememorando algo fundamental na semântica de uma língua: a noção de sinonímia. Os falantes tendem a não ser rebarbativos, repetitivos, circulares, e evitam a sinonímia, embora nem sempre, ao constatarmos que itens como \*[fazi] e [fiz] podem compartilhar uma mesma comunidade de fala.

O aprendiz do português como segunda língua e a criança pronunciam, por vezes, \*[fazi]. Na maioria dos casos, o falante usará uma palavra a partir de seu léxico ao invés de recorrer à morfologia para produzir uma nova palavra com o mesmo significado. Esse fenômeno, “(...) a não ocorrência de uma forma, sujeita a simples existência de uma outra”<sup>4</sup> (ARONOFF 1976, p.43), é chamado de *bloqueio*, e seus efeitos podem ser vistos não somente na flexão, mas também na derivação, em que uma palavra como \*[desmorreu] (formada por [des-] e [-morreu]) será bloqueada por [ressuscitar], que já existe no léxico de um falante. Podemos dizer que o trabalho do bloqueio é tornar \*[desmorreu] inaceitável, porque outra palavra com o mesmo sentido, adequada a um mesmo contexto, é perfeitamente aceitável, quando não existir ainda palavra para bloqueá-la.

---

<sup>3</sup> Does the speaker/hearer somehow check the lexicon to see if a word is there, and only resort to the morphology if there is none?

<sup>4</sup> *Blocking* is the nonoccurrence of one form due to the simple existence of another.

Podemos ver que [desmatou], estruturalmente análoga a \*[desmorreu] é perfeitamente aceita, por não haver supostamente uma outra palavra para bloqueá-la. Segundo Aronoff (2001, p.242), “Os efeitos do bloqueio são também sentidos na sintaxe, em que uma palavra que existe bloqueará, às vezes, um sintagma inteiro, como pioneiramente notou Hoffman (1982)”<sup>5</sup>. Não encontramos, por exemplo, \*[homem da vida], sintagma usado no português num sentido estruturalmente paralelo a [mulher da vida], mas semanticamente inaceitável, por causa da existência de sinônimos como prostituto/garoto de programa. Como o exemplo demonstra, a sinonímia exata, fenômeno raro numa língua natural, é crucial, para a expressão \*[homem da vida] poder de fato ser usada. Ela não pode ser sinônimo de [garoto de programa/prostituto]: será que um *\*homem da vida* pode vir a significar a mesma coisa que um *mulher da vida*?

A noção de bloqueio é uma tentativa sofisticada introduzida por Aronoff (1976) para captar as dificuldades ou até mesmo impossibilidades que os falantes têm em criar novas palavras. Segundo esse autor, repetindo, “a não-ocorrência de um item devido a simples existência de um outro”, que possa ser usado em mesmos contextos, sem perdas semânticas significativas, capta a noção de bloqueio.

Rocha (1998, p.138) propõe quatro tipos de bloqueios para a língua portuguesa: o paradigmático, o heterônimo, o homofônico e o parônimo.

a) Bloqueio paradigmático: quando um item existe em uma língua e está disponível para uso, consagrado por uma tradição, “(...) diremos que muitas vezes não se cria uma determinada palavra pelo simples fato de já haver uma outra correspondente, com o mesmo sentido e/ou função”. (ROCHA, *op. cit.*, p.141). É o caso de bambuzal, cuja *casa lexical*, já preenchida, não fomenta a necessidade de o falante criar algo nos termos (?)*bambual*, (?)*bambuano*, (?)*bambuário*.

b) Bloqueio heterônimo: alguns itens não são criados na língua pela simples existência de outros, que os bloqueiam. Rocha (*op.cit.*, p.141) cita alguns:

terra	(?) terreiro	(bloqueado por agricultor)
unha	(?) unheiro	(bloqueado por manicure)
lua	(?) lueiro	(bloqueado por astrônomo)

c) Bloqueio homofônico: aspectos fonéticos vetando um novo item. Rocha (1998, p.141) elenca, dentre outros:

sala            (?) saleiro            (bloqueado por saleiro – recipiente onde se coloca o sal)

<sup>5</sup> The effects of blocking are also felt in syntax, where an existing word will sometimes block an entire synonymous phrase, as Hoffman (1982) first noted.



cobre (?) cobreiro (bloqueado por ‘cobreiro’ – doença)

d) Bloqueio parônimo: a língua apresenta parônimos que bloqueiam o surgimento de um outro item. Ainda Rocha (*op.cit.*, p. 141):

cabelo (?) cabeleiro (bloqueado por cabeleireiro – base ‘cabeleira’)

vidro (?) vidreiro (bloqueado por vidraceiro – base ‘vidraça’)

A sinonímia de uma forma ou de outra sempre está envolvida na noção de bloqueio. Em português, *férias* não significa uma extensão no plural de *féria* precisamente porque esta palavra se refere, dentre outros sentidos, segundo o Aurélio (*op.cit.*, p.887), a “dia semanal, jornal ou salário de trabalhadores”; enquanto que a primeira significa, dentre outras coisas, “dias em que se suspende o trabalho para descanso”. É tanto que esses dois itens têm entradas diferentes no Aurélio, contrastando com, por exemplo, ‘jornalista’, que possui somente uma entrada no Aurélio, sem ser explicitada a sua extensão no plural.

Segundo Aronoff (2001, p. 240), “A descrição mais convincente de por que o bloqueio ocorre é a de Horn (1984, 1993), baseada em princípios gerais de economia de expressão”<sup>6</sup>.

Não somente questões estruturais, como aponta Rocha (*op.cit.*), acima, estão envolvidas na noção de bloqueio, mas também podemos constatar pressupostos cognitivos envolvidos. Quando um falante “se esquece” temporariamente de uma entrada como [amante], por exemplo, suporíamos que essa pessoa poderia, então, usar, num mesmo contexto, a palavra \*[amador], pois, nesse caso, em sua mente, [amante] estaria bloqueada por questões de memória, nervosismo etc.

Essa suposta falha de bloqueio é especialmente comum em crianças, que cunham novas palavras quase que livremente, pois o seu vocabulário não está tão bem fixado como o dos adultos. Uma criança articulada pode usar palavras como \*[fazeu] e \*[desmorreu] numa conversa, sem hesitar.

O bloqueio é também subjetivo por um outro fator psicológico: a familiaridade ou sua contraparte mais facilmente mensurável, a frequência. Em geral, uma forma irregular usada com mais frequência é recorrente no repertório vocabular de uma pessoa do que a forma regular.

Em certas comunidades de fala, por exemplo, a forma nominal de particípio do verbo *falar* é abundante, convivendo, em contextos bem parecidos, [falo] e [falado], em expressões

---

<sup>6</sup> The most cogent account of why blocking occurs is Horn`s (1984, 1993), based on general principles of economy of expression.

como: “Ela tinha *falo*” e “Ela tinha *falado*”. Geralmente a escola consegue bloquear *falo* como um particípio, fazendo correções constantes da fala dos alunos.

O efeito da frequência pode ser detectado não somente experimentalmente (PINKER and PRICE, 1991), mas também nas *super-regularizações* das crianças, como Bybee and Slobin (1982) demonstraram para os verbos irregulares do inglês. Os modelos de bloqueio mais amplamente aceitos levam a frequência em consideração, que lida com a noção pragmática de processamento. De acordo com esses modelos (MACWHINNEY, 1975; ANSHEN e ARONOFF, 1988; PINKER e PRINCE, 1991), a busca por uma palavra pode ser vista como a corrida entre o léxico mental e a morfologia: ambos operam simultaneamente.

### 3.2 QUANDO A MORFOLOGIA SE APROPRIA DO LÉXICO

Podemos reduzir os modelos morfológicos a uma linguagem formal. A entrada lexical [nacionalização] pode ser vista a partir de sua redução a uma linguagem abstrata, formal, como [[[[X]s al]adj izar]v cão]s. Essa fórmula matemática pode ser lida como:

***[X], base, (s) substantivo; [al], afixo que forma (adj) adjetivo; [izar], afixo que forma (v) verbo; [cão], afixo que forma (s) substantivo.***

Se observarmos mais de perto as palavras que se enquadram nesse modelo, vemos o efeito da morfologia. Vamos limitar o nosso olhar ao padrão geral já mencionado para um deles mais ligeiramente em particular, aquele no qual o adjetivo é de tipo [X]ado (por exemplo, calçado, culpado, apressado). Se o modelo fosse completamente independente do léxico, poderíamos então esperar descobrir que qualquer adjetivo da forma [X]ado poderia servir como base de uma palavra atestada na forma [X]adinho. Levantamos uma ampla lista de palavras do português, compilada dos dicionários Houaiss e Aurélio, contendo mais de 5 mil entradas cada um desses dicionários. Entre essas, aproximadamente 1000 palavras na forma [X]ado mas somente 500 palavras na forma [X]adinho, o que indica que essa regra em particular não é muito produtiva. Dessa forma, não vamos encontrar listadas as seguintes palavras – embora todas elas tenham sentido facilmente deduzível – \*enxadazinha, \*saladinha, \*macarronadinha, \*aladazinha. É bom estarmos atentos à diferença entre não listado e não existente: uma entrada pode não ser listada num dicionário mas isso não significa ou impede que ela não possa existir.

Não podemos, claro, generalizar diretamente a partir de um dicionário para o léxico mental, mas os dados de dicionários desse tipo, que podem ser repetidos para muitos outros modelos nos quais um sufixo é anexado a um outro, sugerem que a produção real de palavras morfológicamente complexas é feita ao se aplicar amplamente regras morfológicas

(adicionando afixos) a palavras-base que existem de fato, e que são estocadas no léxico mental de um falante.

Um outro indicativo de que regras morfológicas operam sobre palavras no léxico é a herança da irregularidade – a morfologia manipula o léxico. O tipo mais comum de irregularidade herdada é a semântica. Entradas complexas geralmente possuem sentidos convencionais que diferem ligeiramente, ou absolutamente (é o caso de *pé-de-molque*) de seu sentido pré-concebido.

Consideremos a palavra *mensurável*. Julgando suas partes, ela deve significar *aquilo que pode ser medido*. O advérbio formado a partir dela, *imensuravelmente*, carrega, dentre outros traços semânticos, os traços *grandemente*, *muito*, em expressões como “Eu me beneficieei imensuravelmente de sua assistência”.

Um outro exemplo é *naturalizar*, que pode significar *tornar natural*, mas que tem um número de sentidos especializados, incluindo “conferir direitos a um cidadão (um estrangeiro)” e “adaptar (uma planta ou animal) a um novo meio-ambiente”. O substantivo *naturalização*, derivado desse verbo, possui derivados nominais dessas duas palavras como significado por ser formado pelo verbo real no léxico com todos os seus sentidos especializados. Ademais, para a maioria das pessoas, o sentido mais familiar do verbo é “conferir os direitos de cidadão a alguém (um estrangeiro)”, o sentido mais saliente do substantivo para a maioria das pessoas, e baseado naquele sentido do verbo.

Para concluir, vimos que a morfologia é distinta do léxico (ao menos se, por léxico, queremos dizer o léxico mental de formas não previsíveis de um falante/ouvinte), e que a morfologia e o léxico são fontes rivais das palavras. A morfologia depende do léxico, entretanto, porque as bases de palavras morfológicamente complexas são normalmente entradas lexicais. A morfologia é o domínio previsível das entradas, enquanto o léxico pode ser concebido como um domínio realizado das entradas lexicais.

### 3.2 ENTRADA LEXICAL E PRODUTIVIDADE MORFOLÓGICA

A produtividade morfológica pode ser definida de maneira informal como a extensão daquilo que um afixo particular provavelmente é, para ser usado na produção de novas palavras na língua. Sob esse ponto de vista, a produtividade é um *continuum* probabilístico que prevê o uso de palavras potenciais. No fim do *continuum* estão os afixos não produtivos

ou completamente mortos (fossilizados), que não são, dessa forma, para ser usados em todas as novas palavras cunhadas.

Um exemplo disso, cunhado por Rocha (1998), é o sufixo [-ebre], considerado um fóssil improdutivo na língua portuguesa, enquanto os sufixos com os quais estamos trabalhando [-dor], [-eiro] e [-nte] são relativamente produtivos em português contemporâneo.

Alguns linguistas tratam a produtividade morfológica como uma noção absoluta – um modelo que pode ser tanto produtivo quanto não produtivo – mas há um bom rol de evidência para a existência e utilidade de casos intermediários, que não é de interesse estrito neste livro. Assim, assumimos provisoriamente que afixos podem se diferenciar continuamente em termos de produtividade, ao invés de recaírem somente em categorias polares de completa produtividade ou completa improdatividade, para as quais alguns linguistas têm restringido a discussão.

À parte considerações quantitativas, há fatores morfológicos qualitativos que são relevantes para a produtividade. Isso pode ser mais bem visto no exame de afixos rivais, afixos que são muito similares em suas condições semânticas e sintáticas. Consideremos uma série de sufixos que formam substantivos a partir de adjetivos no português. Muitos deles são mais ou menos produtivos dentro de um domínio morfológico restrito, é o caso de [-nte] em comparação com [-eiro]. Enquanto o primeiro é mais restrito em termos semânticos, o segundo tem o seu escopo mais ampliado, podendo aludir a uma avaliação negativa da atividade de alguém, como no caso de [bicheiro] e [peniqueiro]. Se voltarmos ao conceito original de produtividade mencionado acima, a extensão para a qual um dado afixo é usado na produção de novas entradas na língua pode ser real como também virtual.

A noção de frequência, que discutimos acima em relação ao bloqueio, está também relacionada à produtividade: quanto menos produtivo um modelo morfológico for, mais frequente será a média dos seus indivíduos membros. Mas a frequência é algo também importante na seleção das bases: um afixo menos produtivo é geralmente encontrado anexado a palavras-base de frequência mais alta que um afixo mais produtivo (ARONOFF, 1982). Isso faz sentido em termos do que sabemos sobre a conexão frequência e reconhecimento lexical: palavras com bases de alta frequência são de mais fácil cognição que palavras com frequência similar mas com bases de baixa frequência (LAUDANNA and BURANI, 1985). É o caso da entrada-sufixo [-eiro] em relação à entrada-sufixo [-nte] com o traço semântico agente.

Se afixos menos produtivos estão em desvantagem, eles então parecem se beneficiar da ajuda fornecida por uma base mais frequente, embora o mecanismo psicolinguístico exato por trás desse padrão não esteja claro.

Alguns estudiosos têm insistido em que o estudo da produtividade morfológica deveria propriamente se confinar ao estudo das palavras que são produzidas sem intencionalidade (SCHULTINK, 1961). Isso exclui inteiramente o estudo da morfologia não produtiva, que se assemelha mais a formas marginais de criação de palavra como a formação por combinações (por exemplo, *bebemorar* como uma mistura de *beber* e *comemorar*) ou acrônimos (como *AIDS*, formado pelas letras iniciais do sintagma *Síndrome da Imuno-deficiência Adquirida*) sendo mais provável que intencional ou notada. Entretanto, os modelos morfológicos seguramente produtivos são os mais acentuados num dado momento sincrônico de uma língua, motivados por fatores funcionais e estruturais, e parecem servir a uma função que nasce de sua acentuada produtividade.

Quando comparamos o conjunto de palavras formadas por meio de um afixo menos produtivo com um conjunto formado por um afixo rival no mesmo meio-ambiente morfológico, geralmente descobrimos que os significados do conjunto formado pelo menos produtivo são menos previsíveis, tornando o conjunto inteiro menos coerente semanticamente. Essa diferença de coerência persiste em palavras cunhadas recentemente: o significado de uma nova palavra formada mediante um afixo menos produtivo será menos assimilável semanticamente, fazendo da morfologia um componente recorrente entre os falantes para predizer o seu significado, enquanto o léxico atuaria de forma mais segura para o falante.

Resumo: Neste capítulo delimitamos diferenças entre morfologia e léxico no conceito de Entrada Lexical. Baseados sobretudo nos trabalhos de Aronoff, vimos como morfologia e léxico se articulam, e o papel que cabe a cada um deles. Deixamos bem claro no decorrer das discussões que acreditamos em modelos de análise linguística que preconizam a existência de um módulo independente da gramática, ou seja, um módulo chamado de Léxico, apesar de estarmos conscientes de que o léxico não vem sozinho na mente dos falantes: ele é manipulado pela sintaxe, como sustentado por modelos teóricos da MD. No capítulo seguinte, vamos explicitar de maneira mais acurada o conceito de Entrada Lexical a partir das noções de unidade lexical e lexema.

#### **4 ENTRADA LEXICAL E DISTINÇÃO ENTRE UNIDADES LEXICAIS E LEXEMAS**

FOCO: Vamos recorrer a pressupostos da Semântica Lexical sobre uma Entrada Lexical e suas relações com as noções de unidades lexicais e lexemas. Os dados serão levantados no romance “O Coronel e o Lobisomem”, de José Cândido de Carvalho, selecionados de maneira aleatória.

O conceito de unidades lexicais, amplamente explorado em estudos semânticos e morfossintáticos para o inglês (HARLEY, 2006) e para o português brasileiro (SANDMAN, 1992; BASÍLIO, 1980; ROCHA, 1998), será aqui retomado tendo em vista as criações lexicais expressivas no romance (CL), de José Cândido de Carvalho. Será necessário, inicialmente, introduzir uma distinção entre dois tipos de entradas lexicais – elementos relevantes para uma descrição em Semântica Lexical: a esses dois tipos, podemos chamar de *unidades lexicais e lexemas*.

Vamos delinear rapidamente, mais abaixo, o que podemos instituir como unidades lexicais e lexemas, mas, neste capítulo, nosso objetivo central será estabelecer bases teóricas consistentes para delimitar as chamadas unidades lexicais, e pressupor que elas compartilham traços semânticos funcionais, enquanto os lexemas compartilham traços semânticos instituídos no Léxico.

A concepção de unidade lexical adotada aqui não difere muito daquela que a literatura lexicográfica convencional tem apresentado. Um dicionário comum caracteriza uma unidade

lexical de três formas distintas, mas intimamente inter-relacionadas: em termos formais, em termos funcionais e em termos semânticos.

- em termos formais, geralmente um verbete de dicionário apresenta sua forma gráfica e sua forma fonológica, e morfológica ou derivacional, apresentando muitas vezes sua estrutura;
- em termos funcionais, entra em cena a função gramatical, em que, para uma unidade lexical, é disponibilizada informação distributiva e categorial, assim como é fixada sua classe; e,
- já em termos semânticos, fixa-se o seu sentido, em que é verificada a sua abrangência em termos polissêmicos e contextuais – geralmente alguns dicionários trazem exemplos retirados de jornais e da literatura para sabermos como, quando e onde uma unidade lexical foi utilizada anteriormente.

Para nos certificarmos da abrangência de um verbete de dicionário, em termos formais, em termos funcionais e em termos semânticos, como acabamos de sugerir, vamos levar em consideração um verbete de um dicionário bilíngue francês/português, português/francês, abaixo:

**GROS, GROSSE (grou, óce) adj.** Grosso, gordo, corpulento; considerável, avultado; tosco, grosseiro; longo, comprido. (*fam.*) Rico, opulento. \_\_\_ TEMPS, temporal. \_\_\_ JEU, jogo forte. \_\_\_ VERT, verde-escuro. \_\_\_ SEL, piada forte. AVOIR LE COUER\_\_\_, estar oprimido. \_\_\_ MARCHAND, negociante opulento. \_\_\_ MOT, palavra ofensiva. UNE GROSSE SOME, uma soma avultada. GROSSE ARTILLERIE, artilharia pesada. *s.m* O grosso, a parte mais grossa de uma coisa; troço, multidão considerável. EM \_\_\_, por atacado. **adv.** Muito. TOUT EM \_\_\_, quando muito. (Dicionário de Francês: Francês/português, Português/Francês, 41ª ed., Editora Globo, S. Burtin-Vinholes, p.245)

Em termos formais, podemos atestar que o dicionário bilíngue investe na forma gráfica do verbete em foco (*gros, grosse*), como também na forma fonológica, nesse caso, uma demonstração de como essa forma é pronunciada em francês (*grou, óce*); em termos de sua forma morfológica ou derivacional, isso pode ser atestado quando o verbete sinaliza o uso do item em sua forma masculina (GROS), como também em sua forma feminina (GROSSE), apresentando muitas vezes sua estrutura: nesses termos, a forma feminina apresenta o morfema de gênero [-s]e, enquanto a masculina está neste gênero por ausência de morfema.

Em relação a sua função gramatical, disponibiliza-se para o verbete informação distributiva e categorial, assim como é fixada sua classe, como podemos verificar pelas informações em negrito: (**adj.**), sinalizando que o item pode ser utilizado numa sentença na

função qualificativa; a informação (*s.m.*) aponta que há ocorrências em que o item é utilizado como um substantivo masculino; e esta outra (*adv.*), neste caso, informa que o item é usado na função circunstancial.

É de senso amplamente aceito que num dicionário busca-se o sentido das palavras, e o dicionário aqui em questão oferece fartas possibilidades de uso, focando a função semântica, em relação a ocorrências de sentido em termos sintagmáticos, como, por exemplo, a ocorrência que segue: [\_\_\_\_ *MARCHAND*, *negociante opulento*]. O espaço vazio [\_\_\_\_] informa que aí deve caber o verbete em análise [*gros*], assim, [*gros marchand*]; não poderia ser colocado (*grosse*) no lugar do espaço vazio pois a palavra *marchand* é masculina.

O verbete em questão aponta possibilidades tanto polissêmicas quanto contextuais, em termos sintagmáticos. Os espaços vazios apontam uma ampla gama de possibilidades de uso em termos sintagmáticos, como acabamos de sugerir.

Vemos ainda o incipiente esboço de uso do verbete em termos discursivos. Isso acontece quando, entre parênteses, aparece a informação uso (*fam*), em que se espera que o verbete em análise seja utilizado em circunstâncias coloquiais – familiar –, em contraste com a possibilidade de uso em termos técnicos, e/ou, em termos mais planejados ou formais na produção linguística.

Vamo-nos apropriar das concepções de Cruse (1992) para estipular um conceito de unidade lexical para o português, tanto paradigmática quanto sintagmaticamente. Nos parágrafos que seguem, vamos delimitar as fronteiras entre os itens lexicais numa sentença, e seus contextos.

Vamos assumir que qualquer sentença bem formada exige uma sequência de itens encadeados conforme cumprimento de regras que encadeiam sentenças, como também regras que *encadeiam* o léxico<sup>7</sup>. Ao organizar conjuntos de unidades lexicais numa sentença, observamos que elas parecem operar sob tutela de uma variedade de investidas gramaticais. Por exemplo:

(1) *Uivava de cortar o coração mais de pedra.* (CL, p.39)

As investidas gramaticais para um item como [uivava] podem ser de tempo, modo, número e pessoa; o argumento<sup>8</sup> à esquerda desse item não aparece, mas sua posição está presente, como demonstrado no esquema: [\_\_\_\_ uivava].

<sup>7</sup> Remetemos os leitores ao capítulo I.

<sup>8</sup> A noção de argumento, tecnicamente usada nesse contexto, está delineada no capítulo *A noção de Entrada Lexical na estrutura argumental*, logo adiante.



Alguém pratica essa ação – em termos semânticos não existe uivo que não seja dado por alguém – mas, apesar de esse alguém não estar presente, o espaço [\_\_\_\_\_] onde ele deve ocorrer (à semelhança dos espaços vazios, antes e depois, no verbete de dicionário acima) está aí por ser preenchido, pois, como fixamos itens lexicais em termos funcionais aqui, pressupomos uma estrutura em que esse item se desloca do léxico (dicionário mental), em estado, digamos, bruto (ainda um lexema da língua), e se projeta numa sentença, flexionado em pessoa, tempo e número. A noção hipotética e intuitiva de estado bruto foi rapidamente delineada numa pequena seção que está no prólogo deste livro, que chamamos de *Limbo Lexical*.

As diversas investidas gramaticais em um item surpreendem por elastecer seu sentido de base. Assim, o passado do item [uivava] pode-se converter em presente [uiva]; a pessoa singular desse item pode ser convertida em plural [uivavam] etc.

Apesar de essas investidas mudarem a abrangência do item, ele não se modifica completamente, pois permanece algo que faz parte intrinsecamente dele. É como se uma variedade de investidas gramaticais não conseguisse cindi-lo, partindo-o em dois ou mais de dois itens. Semanticamente ele permanece clivado, partido ao meio, mas não inteiramente, por podermos reconstituir a sua *personalidade* semântica.

Há uma estrutura gramatical constante para cada item dentro de seus respectivos contextos. Aliás, a concepção clássica, e quase que axiomática, em linguística – a separação entre sentido e gramática – ganha uma justificativa forte: a gramática atua, instituindo e roubando, ampliando e antecipando traços semânticos, mas eles permanecem dando identidade a um item lexical.

A gramática não é capaz de modificar semanticamente um item por inteiro. Se assim o fizesse, o Léxico seria uma lista interminável de itens lexicais, cada um deles reportando a um sentido. Ao contrário, temos uma família semântica (os lexemas) e fazem parte de cada família semântica itens lexicais diversos, modificados por investidas contextuais e gramaticais de variados tipos, compondo um sistema discreto e relativamente estável.

Tomemos as sentenças abaixo como exemplo de uma família desse tipo, e quais podem ser membros possíveis dessa família. Quando estamos dizendo membros dessa família, estamos nos reportando à estrutura da entrada lexical.

(2)*Alisei o queixo, aporrinhado.* (CL, pág. 9)

(3)*Alisei o queixo, aporrinhada.*

(4)*Alisamos o queixo, aporrinhados.*

(5)*Alisamos o queixo, aporrinhadas.*

Temos algo como um lexema tipo-família: [aporrinhado]. Temos investidas gramaticais de gênero e de número que acrescentam, roubam, ampliam, modificam traços desse item. Semanticamente esse item está clivado, partido ao meio, há uma base que balança ao vento dos sabores contextuais, mas há uma estabilidade, mesmo que precária, que o faz pertencer a uma família.

Nos dicionários, por exemplo, não existem tantas entradas assim para ele. Há somente uma entrada tipo texto-colônia<sup>9</sup>, e o verbete remete o leitor para as investidas categoriais e gramaticais possíveis para ele. Como vimos ser feito com o verbete (*gros, grosse*) retirado do dicionário francês/português, português/francês, acima.

É bom notar que as investidas gramaticais e categorial são de vários tipos e agem de formas diferenciadas, ou seja, a investida gramatical para um lexema do tipo como [aporrinhado] é diferente daquela do tipo [uivava].

Vamos estabelecer dois critérios que conjugam as investidas gramaticais e categorial de tal forma a compreender o lado sintagmático de uma unidade lexical, já que estamos operando com o item na estrutura da língua:

- uma unidade lexical é um constituinte semântico de base, e tal constituinte se configura em ao menos um traço semântico estável; e,
- uma unidade lexical é uma palavra da língua arrolada num dicionário mental, ou virtual, sobre a qual agem antes investidas gramaticais e categoriais.

Vamos elaborar cuidadosamente uma explicitação desses critérios, quando focarmos a noção de constituinte semântico, logo abaixo, mas eles servem como uma ilustração preliminar de dois pontos centrais, apontados por Cruse (1992), e adaptados para essa descrição das unidades lexicais para a língua portuguesa que estamos propondo neste trabalho. Esses dois pontos podem ser visualizados a partir das sentenças abaixo:

(6)*Uma ocasião, bem alicerçado de recibos e contas, fui ao ajantarado dos Nogueira disposto a botar tudo no limpo.* (CL, p.220)

(7)*Gato escaldado tem medo de água fria.*

- os afixos em negrito do item [**a** + jantar + **ado**] não são unidades lexicais, embora sejam constituintes semânticos, pois carregam consigo informações gramaticais (dado

---

<sup>9</sup> Um texto-colônia é um termo cunhado por Hoey (2001, p.75) para definir dicionários, glossários, enciclopédias e nomes que aparecem em listas, de um modo geral, como guias de vídeo e almanaques, em que seu discurso é proposto em partes, e “cujas partes componentes não derivam seus significados das sentenças em que estão inseridas”.

que informações gramaticais são informações também semânticas, digamos, de um outro tipo) pois são menores de que uma palavra; para uma palavra é possível manipular contextos sentenciais, para afixos, isso não é possível;

- o item [ajantarado] é uma unidade lexical por estar pronto do ponto de vista funcional, e um constituinte semântico, já que foi manipulado por investidas gramaticais e contextuais (recebeu informações derivacionais [-ado], informações flexionais [-o] etc.); e em,
- [gato escaldado tem medo de água fria] é algo para além de uma unidade lexical, mas podemos dizer que toda a sequência se trata de um constituinte semântico, uma entrada lexical, portanto, dado que esses itens se encadeiam de tal forma que são previsível para a comunidade de fala, e indissociáveis entre si: ocorrendo em contextos semelhantes, e com funções parecidas a uma palavra composta (uma super palavra composta); trata-se de uma metáfora – não estamos dispostos a trabalhar com metáforas neste trabalho, algo que foge ao escopo do tipo de semântica lexical no qual estamos investindo.

Conclusão provisória 1: há constituintes semânticos menores e maiores de que uma unidade lexical, mas uma unidade lexical não pode prescindir de (por ausência e por presença):

- uma constituência semântica, em termos de um constituinte semântico;
- investidas gramaticais, em termos de adequação sintagmática na estrutura de uma língua; e,
- um tamanho: não pode ser menor de que uma palavra, como também não pode ser maior que uma palavra – no caso, ser uma sentença.

Até agora, ainda não nos detivemos acuradamente nas propriedades que fazem parte de um constituinte semântico. Vamos, então, tentar fazer isso nos parágrafos que seguem.

#### 4.1 CONSTITUINTES SEMÂNTICOS

Uma sentença de uma língua natural, adequada a um contexto e nele decodificada seguramente por um contrato semântico<sup>10</sup> entre os usuários de uma língua, é um complexo que resulta da combinação de sentidos entre seus itens componentes.

---

<sup>10</sup> Nada de especial no que propomos com o termo *contrato semântico*. Queremos instituir que uso, produção, construção de sentidos, recepção e acesso a dados da língua, trata-se de que, ao se deparar com uma sentença

Os itens de uma sentença antecipam outros ou podem remeter a outros, num ato contínuo de idas e vindas, tomadas e retomadas. Os sentidos de uma sentença são mais acessíveis do que os sentidos de uma palavra isoladamente, ou mesmo de um sintagma.<sup>11</sup> Isso pode ser afirmado, dada a condição de uma sentença, digamos, completa semanticamente – uma sentença sobrevive linguisticamente, e é eficaz, pois dela fazem parte um contexto, investidas gramaticais, esforços de memória, e esforços pragmáticos dos usuários, e filtrada por um conjunto de regras gramaticais e lexicais atuando tanto na sua formulação quanto na sua concatenação etc. Mas toda sentença exige necessariamente sentidos básicos que fazem parte das unidades lexicais.

Esses sentidos básicos são aqueles plenos que identificam a estabilidade semântica de uma palavra, livre de suas partes gramaticais. Eles vão identificar, de fato, uma sentença; e a forma como eles vão se combinar para produzir um sentido global é indicada pela estrutura sintática da sentença. Vamos utilizar como argumentação a delimitação de um constituinte semântico, tendo o gênero linguístico do português como base para isso, ilustrado na sentença que segue:

**(8) *Era de sem pressa e macia a fala da moça Isabel, modo que ligeiro cativou a minha natureza educada.* (CL, p. 69)**

Não é nosso propósito delimitar um sintagma aqui. Para isso, remeto o leitor para Miotto et alii (1998), Radford (1997), Perini (1996), Raposo (1990), dentre outros. Vamos tomar a sequência [a minha natureza educada] como um sintagma dado. O sentido global desse sintagma surge a partir de duas hipóteses:

- a junção, tipo adição de itens: [a + minha + natureza + educada]; e
- uma relação hierárquica do tipo: [a] não pode aparecer depois de [minha]; [educada] não pode anteceder o sintagma, como em ?[educada a minha natureza], sob pena de mudar completamente o seu sentido e da maneira em que ele se organiza; os itens [a], [minha] e [educada] estão no gênero feminino pois sofrem controle de gênero pelo item-núcleo desse sintagma [natureza].

A estrutura sintática define complexos imediatos, mas as formulações desses complexos imediatos também podem estar contidas na leitura semântica de cada item em si, em outras

---

nova ou um item lexical ‘novo’, “o falante não está diante de elementos novos, mas diante de uma nova combinação de elementos conhecidos” (KATZ e FODOR, pp.81-82), e esse conhecimento pré-construído semanticamente é compartilhado por uma comunidade de fala.

<sup>11</sup> Os sintagmas anômalos em relação aos seus sentidos que estão no capítulo *A noção de Entrada Lexical na estrutura argumental*, ‘salvos’ semanticamente numa sentença, apresentam indícios de que podemos afirmar que os sentidos de uma sentença são mais acessíveis do que os sentidos de uma palavra isoladamente, ou mesmo de um sintagma.

palavras, formuladas potencialmente no léxico. Vejamos como podemos visualizar esses complexos imediatos:

(9)[a], [minha], [natureza], [educada]

(10)[a minha]

(11)[a minha natureza]

(12)[a minha natureza educada]

Dado um complexo imediato do tipo [a minha], podemos deduzir que se espera depois de um item como [a] algo do tipo +feminino, +singular. Podemos definir, então, um constituinte semântico nestes termos: “Qualquer constituinte que faz parte de uma sentença e que carrega um sentido que, combinado a outros sentidos de outros constituintes, institui o sentido mais amplo de uma sentença” (SANTANA, 2005, p.58).

Um constituinte semântico que não pode ser segmentado em unidades menores podemos chamá-lo de *constituinte semântico mínimo*. É o caso de [a], de [minha], de [natureza] e de [educada]. O termo constituinte semântico não está sendo usado para se referir a um sentido de um item isoladamente: mas a um complexo forma-sentido-contexto-gramática; isso quer dizer que é uma forma significativa, no sentido de que atribui ativamente significados nele e para ele, e nele e para a sentença em que ele esteja inserido: no jogo de delimitação de um constituinte semântico, sentido e gramática estão interlaçados e são intercambiáveis.

Podemos compreender isso por observação do *comportamento* de um item como [a]: ele porta informações gramaticais (uma categoria de gênero, por exemplo; e necessariamente um número; assim como pode ser categorizado como artigo, uma letra do alfabeto, por exemplo), como também informações semânticas (gênero feminino).

Vamos delimitar logo abaixo a constituência semântica de uma unidade lexical. Os itens lexicais apresentam extensão de sentido, como firmamos logo acima. Essa extensão de sentido contribui para que a sentença seja interpretada, e a extensão também atua para que a sentença seja interpretada de uma determinada maneira e não de outra, fechando possibilidades semânticas.

Por exemplo, a investida gramatical que incide sobre um item como [uivava], que o faz aparecer no passado, incide também sobre toda a sentença. Assim, toda a sentença está no passado, a partir da leitura de um item em particular. Essa delimitação é necessária para separarmos dois aspectos da constituência semântica:

- há uma investida em termos forma-contexto-gramática que caracteriza uma sentença em relação a si mesma e em relação a uma outra sentença; e,
- uma entrada lexical carrega algum sentido constante de um contexto sentencial/pragmático a outro.

Ao carregar essa constante semântica, uma entrada lexical se constitui como semântica. Em outras palavras, um constituinte semântico porta uma certa estabilidade de interpretação. É como se ele se expandisse como se fosse uma onda no mar: vai e volta e continua a mesma. É como se a sua expansão não o levasse a lugar algum e o levasse a vários lugares ao mesmo tempo.

Se fôssemos recorrer à morfologia derivacional para tornar esse argumento mais explícito, poderíamos utilizar as palavras de Basílio (1980, p.17):

Uma proposta para uma teoria do léxico (...) nos permite: a) captar generalizações entre entradas lexicais, mesmo quando estas não correspondem a processos produtivos de formação de palavras; b) resolver um problema tradicional na morfologia derivacional, ou seja, a análise da estrutura de palavras cujas bases não são itens lexicais que ocorrem como formas livres dentro da língua; c) explicar por que regras de formação de palavras podem operar em bases que não são formas livres na língua; e d) especificar as condições sobre as quais esse fenômeno pode ter lugar.

O item (a) pode nos ajudar a aprofundar a caracterização de um constituinte semântico. Ao propor “captar as generalizações entre entradas lexicais”, a autora estabelece bases para que a constante semântica se apresente como regras abstratas de permanência de sentido, como também significados funcionais em potencial de um item que seja um constituinte ou que faça parte de um constituinte. Há uma conexão entre as contribuições semânticas generalizadas para um item, como o que ocorre nas sentenças abaixo:

(13) *Debelada a ciumeira, de novo nas boas graças de dona Esmeraldina, fui dormir em paz.* (CL, p.235).

(14) *O ciúme corrói a alma de todos nós.*

(15) *Enciumado, Adalberto não se continha...*

e a relação que esse item mantém com outros numa sentença.

Os falantes reconhecem uma relação, algo como uma constante semântica entre [ciumeira], [ciúme] e [enciumado], os itens se expandem (um tipo de derivação, por exemplo, entre [ciumeira] e [enciumado]): um item de base [ciúme] que se *elastece*, mas não perde o *controle* sobre si mesmo, uma espécie de *controle* que é semântico.

Unidades lexicais são itens que reúnem uma complexidade em termos forma-sentido (CRUSE, 1992, p.49) com propriedades semânticas relativamente estáveis e discretas (apontamos acima essas propriedades, mas, mais adiante, vamos explorar mais detidamente o

que elas significam), que estabelecem tanto relações sígnicas de variadas matizes, tais como, antonímia (alto/baixo) e hiponímia (terra/planeta), por exemplo, como também sintagmáticas em contextos variados.

Podemos nos arriscar a dizer que, sem uma oposição, sempre presente e sempre evocada, entre sintagma e paradigma, é hipoteticamente imprevisível e semanticamente inviável estabelecer tais relações exaustivamente, ou admitir que elas sejam, no mínimo, aceitáveis.

Uma unidade lexical particular expressa sua identidade semântica em ambientes contextuais de variados tipos, mas o que nos interessa é redesenhar semanticamente o contexto discursivo-literário do qual emergem as unidades do romance CL, como também de dados aleatórios cunhados de nossa experiência como falante da língua.

Já os lexemas são também uma unidade, porém, listados no léxico, ou numa espécie de dicionário mental de uma língua. Ao constituirmos unidades lexicais e lexemas nesses termos, podemos admitir:

- a) Um item é, ao mesmo tempo, uma unidade lexical e um lexema.

Tomemos como exemplo a sentença que segue, recortada de CL:

(16) Espicacei o orgulho do **galista**. (CL, p.69)

O item em negrito (**galista**) funciona tanto como um lexema, pois está previsivelmente listado no dicionário mental dos falantes nativos da língua portuguesa, que reconhecem partes menores desse item, tipo (*gal* + *ista*), e sabem jogar com essas partes, tanto mental como produtivamente, ao reconhecer e/ou produzir itens outros, do tipo (*golp* + *ista*) ou (*gal* + *inha*); como também esse item em negrito (**galista**) funciona como uma unidade lexical, por manter relações semânticas aparentemente imprevisíveis, arroladas significativamente durante a leitura do romance, e ter vestido uma roupagem gramatical, semântica e funcional na sentença. Em outras palavras, [**galista**] é uma criação lexical disponível no romance em estudo, criado especialmente para fazer parte desse micro-universo literário.

- b) Apesar de um item participar tanto do quadro teórico da unidade lexical quanto do quadro teórico de um lexema, podemos admitir, para o tipo de análise que nos interessa, que só como unidade lexical, a partir dos critérios que estamos instituindo, é possível arrolar traços semânticos interpretativos relevantes, e mais, quando unidade lexical ele é um lexema, mas como lexema é necessário

um contexto, potencialmente funcional e/ou sentencial, para que se transmute em unidade lexical.

Suponhamos o seguinte: o que faríamos se alguém nos perguntasse o que significa [*galista*]? Poderíamos até encontrar mentalmente possibilidades de respostas, mas não teríamos um contexto em comum em que pudéssemos incluir um sentido compartilhado por uma comunidade de fala.

No romance, contudo, temos um contexto a compartilhar, um contexto, nesse caso, textual-discursivo, em outros termos, um contexto literário. Isso significa que, nesse contexto, e talvez só para ele, poderíamos constituir um sentido para a palavra **galista**.

Pode parecer estranho ou redundante, talvez até desnecessário, apresentar ou delimitar dois tipos de itens em termos semânticos: lexemas e unidades lexicais. Como há, contudo, diferentes funções, que impõem diferentes contrastes para delimitar esses itens nesses termos, defendemos essa delimitação, entretanto.

Os itens lexicais compartilham, portanto, *traços semânticos* e *funcionais*. Só que esses traços funcionais são, de alguma maneira, semânticos também. Vamos estabelecer, por exemplo, a noção de oposição para lidar melhor com esses traços.

Haveria necessariamente um jogo de a) *fricções*, e b) *oposições* que só funcionalmente pode ser demonstrado, e, dificilmente ou improvavelmente, poderia ser *visível* no léxico. Tome-se como ponto de partida a sentença que segue:

(17) - *Aquilo é tábua de passar roupa. Moça para o primo tem que tercoxão fornido, capaz de agüentar os repuxos.* (CL, p. 17)

a) Em termos lexicais, ou seja, fazendo parte de um dicionário mental, [*tábua de passar roupa*] é uma sequência lexicológica de [nome/tábua; preposição/de; verbo/passar; nome/roupa], e cada item desses é um lexema. Em termos funcionais, contudo, *toda a sequência* funciona como qualificativo nessa sentença. Da *fricção* desses itens podemos designar algo como sequência qualificativa.

Veja-se a sequência abaixo para a demonstração simples do jogo de oposições:

(18) *Uma prima, filha do sepultado tio Tomé de Azeredo, ficou toda ensabonetada para o meu lado.* (CL, p.16)

b) A beleza do item *ensabonetada* pode significar o seu oposto em termos funcionais. Se antepuséssemos uma partícula negativa como [não], por exemplo, ao lado desse item, ele passaria a significar o seu contrário.



Em termos semânticos, um lexema existe em estado de dicionário, mas sobrevive como unidade lexical em termos funcionais, podendo, assim, se transformar, inclusive, em seu oposto quando contaminado semanticamente pelos itens ao seu redor.

Um lexema, por um lado, pode ser associado a um número indefinido de sentidos; quando passa a unidade lexical, passa a ter esse número indefinido de sentidos delimitado em relação ao seu estado de dicionário. A semântica dos itens é relativa e precariamente estável, é bom não perder isso de vista toda vez que nos reportarmos a algo como *estabilidade semântica*, ou seja, a precariedade dessa estabilidade. Nesses termos, as possibilidades de leituras podem ser indefinidas, mas, ao mesmo tempo, estáveis.

#### 4.2 CRIATIVIDADE GRAMATICAL E LÉXICO

A criatividade do falante atuando com sua gramática numa língua tem sido exaustivamente demonstrada, desde as fundações gerativistas do final da década de cinquenta: um número ilimitado de sentenças pode ser produzido a partir de um conjunto limitado de elementos ordenados por regras combinatórias.

Vamos tecer algumas considerações sobre a gramática gerativa, pois nos interessa sobremaneira investir em algo como *universais semânticos* e suas conexões com uma proposta de semântica lexical, seus desdobramentos na delimitação de unidades lexicais, como também antecipar a capacidade combinatória dos falantes em projetar lexemas numa sentença para que eles se transformem em entradas lexicais, e esse mecanismo é universal.

A semântica é o ramo da linguística que se devota a investigar o sentido das palavras, frases e sentenças, assim como a interpretação das expressões de um sistema linguístico. Essa talvez seja a constatação mais antiga e recorrente, e consagrada de longas datas nos estudos linguísticos a partir de variados recortes que lidam com a noção de significado.

Não tentaremos fazer aqui um exame compreensivo dos muitos diferentes recortes das pesquisas recentes da linguística sobre semântica, mas, ao invés disso, selecionar, como já apontamos acima, a Semântica Lexical, e detalhá-la ao máximo que pudermos, tanto em termos descritivos quanto em termos analíticos.

Muitos dos conceitos e técnicas de análise que introduzimos têm suas origens na lógica e na filosofia da linguagem. Vamos aplicá-las ao estudo do léxico e retomar algumas idéias propostas por Miller (1982), Chierchia (2001) e outros ao assumir algo como *universais semânticos*, como pontuamos acima.

Quando dizemos que o nosso foco é na semântica como um ramo da linguística, estamos adotando uma concepção particular de métodos e objetivos da pesquisa em linguística. Essa concepção tem raiz no paradigma gerativo. Desse paradigma, vamos pontuar, para o que nos interessa, três idéias-chave:

- a) A gramática como um mecanismo de regras abstratas.
- b) A linguística pode se estabelecer como uma investigação cognitivista do funcionamento da linguagem.
- c) A linguística não pode apenas se reservar a descrever os dados externos, como uma língua é falada e utilizada na sociedade, – ela precisa analisar não somente empírica mas teoricamente o funcionamento de uma língua natural.

A primeira idéia, a de que a gramática de uma língua pode ser vista como um conjunto abstrato de mecanismos, sistema de regras, e princípios, que servem para caracterizar formalmente várias propriedades da boa formação de sentenças de uma língua, não é tão recente assim, e não tem origem necessariamente nos trabalhos de Chomsky.

No *Cours de Linguistique Generale*, Saussure já propunha essa ideia ao separar a *langue* da *parole*. A gramática, nesse sentido, gera a linguagem. O que era novo era a intuição de Chomsky em pretender que isso seja algo universal, e aplicá-lo às línguas naturais – o tipo que aprendemos a falar e com o qual nos entendemos desde tenra infância – e que poderia também gerar esse sistema formal das línguas naturais.

Em um certo sentido, quando os linguistas adotaram esse ponto de vista, adotaram também a idéia de que uma linguística teórica é um ramo da matemática aplicada, da física teórica contemporânea e da química.

Alguns linguistas gerativos, no entanto, poderiam estar completamente confortáveis com uma caracterização dessa disciplina nesses termos. Uma grande razão subjaz na segunda ideia-chave que Chomsky introduziu, e aqui a mencionamos: a de que as gramáticas gerativas são realidades psicológicas e cognitivas no sentido de que constituem modelos acurados de um conhecimento implícito que repousa na produção e na interpretação das sentenças empreendidas pelos falantes nativos.

O próprio Chomsky nunca pontuou que a linguística fosse um campo da matemática, e sempre a descrevia como um ramo da psicologia cognitiva, e aplicava os modelos matemáticos ao estudo dos fenômenos cognitivos do conhecimento da língua, que muitos linguistas reconhecem como sendo seus objetos de estudo.

Mas, como linguistas, nosso foco está na formulação de um modelo de sistemas cognitivos cujas operações, de algum modo, expliquem os fenômenos linguísticos. A linguística é uma ciência empírica, e, nesses termos, é como a física ou a matemática pura.

A terceira ideia, que vamos delinear, oriunda do paradigma gerativo, é uma interrelação entre essas duas anteriores: a linguística não pode se limitar à documentação do que é dito e de como isso é interpretado – nossa *performance* como falantes e ouvintes –, ainda que os físicos possam limitar seus objetos de pesquisa à documentação e às leituras métricas dos fenômenos físicos diretamente observáveis.

O conhecimento linguístico que buscamos modelar, a *competência* do falante, pode ser distinto de seu comportamento linguístico observável. Tanto a linguística quanto a física pressupõem entidades teóricas abstratas que nos ajudam a explicar os fenômenos observados e prever para além das observações, sob condições específicas.

A distinção entre *competência* e *performance* tem sido muitas vezes saturada e frequentemente mal interpretada. Queremos enfatizar que não estamos exigindo-a a fim de pretender que linguistas devam ignorar a *performance* – as observações de como as pessoas usam a língua são absolutamente relevantes para uma teoria linguística. Pelo contrário, a distinção é importante precisamente porque as observações do comportamento linguístico ocorrendo naturalmente são tipos críticos de dados contra os quais os linguistas gerativos testam suas teorias.

Esses não são, entretanto, somente os dados disponíveis. Por exemplo, um linguista, com frequência, solicita aos falantes nativos (algumas vezes ele próprio) julgamentos intuitivos de como certas sequências de palavras numa dada língua constituem uma sentença gramaticalmente bem formada dessa língua.

Tais julgamentos são também dados para uma análise. Pois bem, se o Léxico está sendo instituído como dados cognitivos da mente humana, como dissemos – um léxico mental – podemos deduzir, assim, que ele pode ser submetido a alguns pressupostos teóricos instituídos pelo arcabouço teórico da gramática gerativa, dentre esse arcabouço, destacamos:

- o léxico é gerativo, por ser recursivo – uma palavra não morre, renasce em outra;
- o léxico é criativo;
- há universais semânticos aplicados ao léxico, e vamos tentar dizer isso em língua portuguesa, pois isso já foi dito em língua inglesa, por Chierchia (2004), Miller et alii

(1982), como também pelo próprio Chomsky (1957, 1970, 1980), quando problematiza a gramática universal.

Vamos tecer alguns comentários logo abaixo a respeito da criatividade lexical, tendo em vista os trabalhos de Basílio (1980) para o português e os trabalhos de Aronoff (1998) para o inglês, amplamente também já resenhados na literatura linguística.

Mas, o que nos interessa agora é investir em ideias sobre o que significam universais semânticos e questionar se isso é possível, e válido, para as línguas naturais, como também para o léxico, assim como a proposta de gramática universal, ainda que polêmica, pode ser aplicada para as entradas lexicais. Vamos começar apresentando a proposta de Chierchia (*op.cit.*).

A gramática gerativa é uma ampla tentativa de caracterizar a capacidade constitutiva das línguas humanas, isso já é amplamente aceito. Dentro de seu grande projeto, que já dura cinquenta anos, estão a investigação da criatividade humana para a linguagem, a instituição de uma faculdade de linguagem humana, as estruturas cerebrais humanas como depositárias dessa capacidade.

Nos anos setenta e começo dos oitenta, novos argumentos ampliaram a perspectiva investigativa da gramática gerativa: a noção de regência e ligação da estrutura sintática, os princípios e parâmetros e algo como a variação paramétrica, e hoje em dia faz escola o recorte minimalista. Assim, pressupomos algumas assertivas, e levantamos uma questão relevante. As assertivas:

- a) A diferença entre as línguas está também no léxico de cada uma delas. Por exemplo, diferenças fonológicas entre [*enfant – guy – garoto*] possibilitariam marcar cada uma dessas entradas lexicais como de línguas diferentes.
- b) O léxico português falado em Portugal e o léxico português falado no Brasil indicam variações entre o português falado aqui e o falado na Europa, mas não é capaz de separar essas duas línguas como sendo línguas distintas.

A questão relevante:

- c) A maneira de significar<sup>12</sup> as entradas lexicais seria diferente de língua para língua ou poderíamos pressupor estratégias universais de significação?

Para enfrentar essa questão, propomos uma tipologia semântica análoga àquela tipologia paramétrica amplamente delimitada nos estudos em gramática gerativa. Propomos universais semânticos, e conceituamos esses universais como estratégias de significar,

---

<sup>12</sup> Há estratégias de significação que percorrem micro unidades lexicais, se espalham por variados níveis, que vão dos morfemas ao discurso.

comuns a todas as línguas. Por exemplo, a ambiguidade e/ou a desambiguação tanto sintática como lexical talvez sejam comuns a todas as línguas.

Segundo Chierchia (2004, p.11), “No final dos anos sessenta e começo dos anos setenta, (houve) algumas tentativas interessantes em estabelecer princípios semânticos universais que culminaram nos assim chamados semanticistas gerativos”.

Ampliando os pressupostos da gramática gerativa para a investigação em semântica, esses linguistas objetivavam que a semântica se estabelecesse como um módulo da gramática com níveis abstratos de interpretação por intermédio de um estoque disponível de palavras ou conceitos básicos concebidos à semelhança do nível sintático. Haveria, então, uma sintaxe de base semântica envolvendo algumas categorias e regras que as combinavam – regras semânticas. Essas idéias se popularizam nos anos noventa nos seguintes termos: “o léxico determina a sintaxe”, (PINKER, 1994, p.48).

Como pontuamos neste trabalho, a noção de argumentos interno e externo dos verbos é uma boa base teórica para justificar o caráter de uma semântica universal focada nos entradas lexicais. Por exemplo:

(19) (\_\_\_\_ X \_\_\_\_)

Na sequência, (\_\_\_\_ X \_\_\_\_) em que (X) seja uma entrada lexical da categoria verbo e os espaços à direita e à esquerda sejam preenchidos por argumentos, informações semânticas contidas nesse verbo (X) selecionariam argumentos, e o tipo de argumentos que devam preencher esses espaços, como exemplificaremos melhor páginas adiante. O problema é o seguinte: a estrutura categorial de uma língua determina que tipo de entrada lexical deve preencher os seus espaços, ou os itens determinariam as categorias e os espaços da estrutura dessa língua? Em outras palavras, o léxico determina a sintaxe ou a sintaxe é determinativa para o léxico? Haveria informações semânticas universais contidas nas palavras que indicariam onde elas devam aparecer nas sentenças?

Vamos aproveitar e expor duas questões relacionadas a isso, feitas por Chierchia (*op.cit*):

a) Tais noções são parte da gramática universal ou há uma outra forma de pensá-las?

b) Elas estão mais conectadas a fenômenos cognitivos ou mais a fenômenos linguísticos?

Essas questões são enfrentadas de maneira parcial, mas já há alguns pressupostos desenvolvidos, a que Chierchia (2004) denomina como:

a) coisificação; e,

b) representação;

desdobrados em dois processos, um de natureza mental:

a) a significatividade cognitiva;

e um outro de natureza linguística:

b) a significatividade informacional.

Vamos por partes. A coisificação (ou reificação) e a representação. Os sentidos se manifestam basicamente pela ligação entre as formas linguísticas e as coisas, sobre as quais falamos, pelas quais nos entendemos e com as quais nos comunicamos.

Essa reificação da língua é algo tão comum que pouco paramos para nos dar conta de que essa é a realidade linguística mais básica em termos semânticos. Mas o fato de que as entradas lexicais portam sentidos que nos permitem usá-las para expressar mensagens e veicular informações não é algo tão óbvio assim em termos de uma investigação científica, já que estamos o tempo todo entre uma *corda bamba* semântica: as coisas são denominadas pelas palavras com as quais as denominamos ou as palavras representam as coisas?

Em outros termos, as palavras são as coisas que portam nomes ou as coisas são representadas linguisticamente por palavras? Se não nos fizermos essas perguntas constantemente a questão do significado perderia sentido.

Por um lado, daríamos como dado que as palavras são as coisas, e perderíamos a noção de que as denominações das coisas surgem historicamente. Por outro, se concebêssemos os sentidos das coisas como ideológicos o tempo todo, perderíamos a noção da estabilidade semântica das entradas lexicais, e reduziríamos os processos mentais a algo como uma consciência linguística, como se tivéssemos consciência o tempo todo do que dizemos e de como significamos. Uma língua escapa aos processos da consciência, ao fazer parte da natureza humana que chega até a ser apontada como um instinto (PINKER, 1994).

Alguns teóricos semanticistas, porém, têm pensado que a coisificação não faz parte do domínio da semântica. Eles focalizam, por exemplo, as estruturas cognitivas que representam o sentido, tomando como significatividade fundamental da língua as relações entre expressões linguísticas, que eles chamam de *representações semânticas*.

A significatividade de uma língua, que propomos aqui como a capacidade de atribuir sentido, inerente à mente humana, pode ser pensada como envolvendo tanto os componentes representação quanto de coisificação.

Diferentes teóricos (CHIERCHIA, op.cit.; JAKENDORFF, 1976; FODOR, 1972) enfatizam a necessidade de interrelação e de conexão entre esses dois aspectos da língua, mas

separam dois componentes que se juntam para refletir melhor sobre esses dois aspectos: a significatividade informacional e a significatividade cognitiva.

A significatividade informacional quer dar conta do aspecto de cosificação de sentidos: por exemplo, que uma pedra é dura, é algo axiomático e universal, apesar de haver pedras que são moles, maleáveis etc. Esse aspecto quer dar conta também das relações entre a língua e as coisas que falamos sobre o mundo.

O componente informacional procura delinear as entradas que fazem parte de uma comunidade linguística e sublinhar o papel das mensagens em termos de um valor de verdade. É possível provar que uma pedra é dura, e isso valer como algo dado para qualquer ser humano numa comunidade de fala.

Já o componente cognitivo envolve as ligações entre a língua e os processos da mente humana, os construtos mentais, que de alguma forma representam ou codificam o conhecimento semântico dos falantes.

O componente significatividade cognitiva quer investir no aparato mental, e não se preocupa com questões que envolvam a realização numa comunidade de fala da comunicação humana. Podemos inferir que tais processos são capazes de cumprir uma *agenda* teórica para universais semânticos. Em resumo: seria possível em termos de universais semânticos admitirmos: os aspectos coisificação e representação, delineados por processos como significatividade informacional e significatividade cognitiva.

Vamos investir um pouco mais no que significa a chamada significatividade informacional nos parágrafos abaixo.

#### 4.3 A SIGNIFICATIVIDADE INFORMACIONAL E COGNITIVA

A significatividade informacional nos permite dizer coisas sobre o mundo, veicular mensagens de uma pessoa para outra, sobre nós mesmos, e sobre as coisas que nos circulam, ressaltando, sobretudo, formas e fatos das coisas.

Quais seriam as propriedades da língua e seus usos abarcando essas formas e fatos? Como uma língua nos permite saber das coisas e nos servir como guias no mundo e nos permitir também apreender o que os outros nos falam, o que ouvimos, sentimos, cheiramos, comemos, sem haver duplicidade em nossas formas perceptuais de sentido? Em outras palavras, quando pedimos a alguém que pegue o *vaso amarelo*, por que alguém pega o vaso amarelo e não o azul (ressalvados os casos de patologia ocular ou cerebral)?

Não há uma relação concreta entre a cor azul e o azul de um objeto, sabemos que isso é uma convenção linguística, faz parte do acordo semântico de uma comunidade de fala. A significatividade informacional não requer necessariamente que a língua se refira ao mundo de maneira predeterminada pela estrutura física do nosso meio-ambiente, e isso é um aprendizado bem cedo na vida de quem adquire uma língua.

Mas podemos constatar o seguinte: se designamos o céu de azul, e se mais tarde encontramos um objeto nunca visto antes, e esse objeto é azul, é evidente que existe uma predeterminação no meio ambiente físico para isso, filtrado pelo domínio do léxico. Quando Gagarin disse certa vez “A Terra é azul”, ele nunca tinha visto a Terra daquele ângulo, mas identificou o azul com o qual já estava acostumado, o qual provavelmente aprendeu por convenção.

A significatividade informacional, portanto, dá-se por convenção, num primeiro momento, mas se amplia para, de fato, predeterminar as coisas num meio ambiente humano. Em primeira instância, segundo Chierchia (2004, p.13)

não se exige apenas que a informação do meio-ambiente seja simplesmente registrada ou percebida sem ativar um *input* da percepção e do pensamento da mente humana. Ainda, isso provavelmente exija uma correspondência regular e sistemática entre a língua e o meio-ambiente que compartilhamos, que é publicamente acessível de muitas maneiras diferentes pelas mentes humanas.

Assim, a semântica da significatividade informacional inclui não somente a relação entre fenômenos concretos ou físicos, mas também mentais ou abstratos. Há questões filosoficamente mais profundas que podem surgir da configuração de tipos diferentes de fenômenos, mas o fato relevante é que, para a semântica, a informação que temos sobre as coisas nos permite falar sobre elas como se fossem dogmas. Permite que as mesmas coisas sejam azuis para todas as pessoas numa comunidade de fala, e ao mesmo tempo.

Em outros recortes semânticos, a significatividade informacional compõe um quadro teórico de algo chamado teorias da referência. Uma semântica das condições de verdade é um tipo particular de teoria da referência. Remetemos o leitor para o trabalho de Katz (1996).

Grande parte das questões de significatividade das línguas tem sido tratada sob uma perspectiva cognitiva. A idéia básica e geral é que temos maneiras de representar mentalmente, e isso significa assumir que o que temos a dizer e a ouvir dos outros fazemos por padrões mentais semelhantes aos que os outros utilizam para nos dizer algo e nos entender também. Há algo na mente humana que nos indica que uma frase como:

(20)[*O carro está limpo*]

pode ser comum para todas as pessoas diante de uma realidade contextual compartilhada.



Aprendemos nos parágrafos anteriores que há uma convenção que nos indica o que é uma coisa chamada [carro] e um processo chamado [limpeza], mas relacionar essa coisa a um processo ultrapassa as noções da convenção e se desloca para algo na mente que nos permite juntar [carro] e [limpeza] na sentença [*o carro está limpo*].

Em outros termos, estamos centrando o foco aqui na capacidade das pessoas em fazer julgamentos sobre as coisas e os processos, no entrelaçamento entre coisas, relações, fatores linguísticos e não linguísticos. Como pontuamos constantemente aqui, conhecer o léxico simplesmente não basta. Conhecê-lo significa relacioná-lo de tal forma que ele seja *apenas* o ponto de partida para se construir sentenças numa língua. O sucesso da comunicação se abre em duas frentes: fazer *links* entre o mundo das palavras e o mundo das situações, como também saber que estamos constantemente em processo de representação.

Ter-se intuições sobre as propriedades semânticas das entradas e expressões não basta, pois podemos possuir essas intuições mas elas talvez não façam parte das mentes das outras pessoas: as mesmas entradas que estão na minha mente podem não fazer parte da mente de uma outra pessoa.

Contudo, muito seguramente, os mesmos processos de relacionar palavras e cumprir regras podem ser comuns a todas as pessoas. Por exemplo, talvez não saibamos o que significa uma palavra como [linguística], mas todos os falantes nativos sabemos que uma frase como [o carro está limpo] faz parte da língua portuguesa.

Com certa segurança é possível dizer que ninguém domina ou conhece todos os lexemas de uma língua, mas que todos os falantes nativos do português sabem como lidar com a estrutura da língua portuguesa, e isso é um processo mental, como se pudéssemos lidar com uma significatividade cognitiva que é parte inerente à mente humana. Ou seja, pensar o sentido nesses termos é pensar como os conteúdos são mentalmente representados. Uma dada configuração de símbolos só teria sentido para nós se conseguíssemos penetrar no seu conteúdo – e tê-lo para nós como se ele nos pertencesse – e isso envolve representações mentais, e tais representações são internalizadas e cruciais na mediação entre os símbolos e os conteúdos. A questão que assombra as teorias psico cognitivas é mapear ou deduzir essas propriedades da mente, o que foge ao escopo de nossas investigações sobre entrada lexical.

Num capítulo intitulado *Mentalês*, Steven Pinker (1994, p.61) tenta fazer isso e levanta uma questão que nos é cara aqui:

O pensamento depende das palavras? As pessoas pensam literalmente em inglês, cherokee, kivunjo (...)? Ou será que nossos pensamentos se expressam por algum

meio de comunicação silencioso do cérebro – uma linguagem do pensamento, ou ‘mentais’ – e só se revestem de palavras quando temos de comunicá-lo a um ouvinte? Não há questão mais central para compreender o instinto da linguagem.

E por extensão, podemos propor essa rede de relações que estamos discutindo aqui, qual seja, uma semântica das coisificações e uma semântica das representações. Essa é uma discussão que faz história em linguística, e tem repercussão direta no conceito de entrada lexical que estamos fundamentando em nosso trabalho, ao retornamos à proposta do determinismo linguístico: a linguagem determina o pensamento ou o pensamento determina a linguagem? Em outras palavras, atribuímos sentido às entradas ou os sentidos atribuídos a elas já nos chegam relativamente prontos e os utilizamos conforme as convenções?

Pinker radicaliza e tenta desmontar o determinismo linguístico, mas não consegue nos fazer crer que exista algo a que ele atribui o nome de *mentais* – estamos sempre numa corda bamba semântica, como já apontamos acima: determinamos a linguagem e somos determinados por ela, ou seja, as coisas podem ser representadas como também podem ser coisificadas por intermédio das entradas.

O que queremos estabelecer aqui é a relação indissociável entre uma teoria das representações das palavras e uma teoria da significatividade informacional que conjugam pensamento, coisificação, cognição e representação.

Segundo Austin (1962, p.30), “dizer algo pode significar fazer algo”. Isso quer dizer que, ao mesmo tempo em que pronunciamos frases e sentenças, em verdade estamos fazendo coisas ao mesmo tempo em que estamos dizendo coisas. Nesse caso, separar representação de coisificação é algo extremamente problemático. Pontuamos, entretanto, que essas coisas interessam sobremaneira à semântica lexical, simplesmente porque o léxico permeia todo esse debate e é pelo léxico que esse debate é possível.

A criatividade lexical é similar à criatividade gramatical. De um número limitado de entradas do dicionário mental, o falante pode produzir um número ilimitado de sentidos em termos funcionais quando faz a passagem do lexema à unidade lexical. Portanto, a unidade do léxico mental, que varia de falante para falante, mas pode ser quantificado em termos de entradas lexicais, submetidas a regras combinatórias, pode permitir a produção de um número possivelmente ilimitado de novas unidades, não especificamente quantificáveis.

Quem poderia admitir que [*tábua de passar roupa*] se transformasse em um qualificativo, ‘de uma hora para outra’? Acima, deduzimos que os itens são relativamente estáveis e discretos. Eles são estáveis em termos de lexema, num dicionário mental; e discretos em termos de produção semântica máxima a partir de um número reduzido, em

termos quantificáveis, de entradas lexicais. As entradas em um dicionário podem ser até contadas. Isso serve até de propaganda para um dicionário: como “Essa edição do Aurélio vem com sessenta mil verbetes”.

Certa vez, já ouvi falar em pesquisas inúteis do tipo: um analfabeto possui um vocabulário reduzido a X entradas, enquanto uma pessoa altamente escolarizada pode conter um número maior de entradas, como se o léxico mental fosse reduzido a um cálculo matemático.

Conclusão provisória 2: As entradas lexicais, construto hipotético submetido a verificações em termos mentalistas e em termos públicos (visualizados em dicionários como o Aurélio e o Houaiss), se associam a possibilidades de concatenação do léxico para fins cognitivos e funcionais.

#### 4.4 DELIMITANDO LEXEMAS

As problematizações apresentadas que fazem parte desta seção são levemente inspiradas em Cruse (1992).

Uma das características mais marcantes de uma língua é o fato de que é possível construir um número ilimitado de sentenças e sentidos a partir de recursos e fontes finitas, delimitadas por critérios morfológicos (morfemas), fonológicos (fonemas) e sintáticos (ordem etc). É amplamente aceito isso para a gramática de uma língua, e a história da teoria gerativa é a história para argumentar que seja possível.

Vamos argumentar que isso é possível para a semântica também, mais especificamente, para o domínio lexical tendo como foco a noção de entradas lexicais em suas relações com o léxico. Há uma multiplicidade de propostas de estudo do léxico: a lexicografia, a lexicologia, a terminologia, dentre elas. A nossa proposta é o estudo baseado na Semântica Lexical, como já pontuado algumas vezes aqui no decorrer deste trabalho.

Há um conjunto de elementos finitos, enumeráveis, dos quais uma proposta em semântica lexical quer dar conta. O número de unidades finitas apropriadas para o nosso propósito aqui, queremos chamar de *lexemas*: um dicionário contém, entre outros elementos, uma lista alfabética de palavras de uma língua natural. Sobre palavras, há investidas em termos morfológicos (bases, afixos, morfemas, número, gênero etc), em termos sintáticos (argumentos de um verbo, núcleos de sujeito, predicado etc).

Quando, porém, propomos investir nas palavras em termos semânticos, devemos delimitar o nosso objeto com características e um jargão apropriado: em termos semânticos, uma palavra assume a condição de *lexema*. Antes de delinear as propriedades de um lexema, contudo, devemos investir em conceitos como dicionário e dicionário mental (já aludidos no começo deste capítulo II), já que podemos antecipar que um dicionário é um arquivo onde estão depositados os lexemas de uma língua. Tarefa para os parágrafos que seguem.

Propostas para conceituar ou delinear o que é um dicionário de uma língua já estão amplamente disponíveis para um mercado linguístico, como também já possuem um campo específico em linguística: a lexicografia.

Segundo Biderman (1996, p.131): “Os dicionários constituem uma organização sistemática do léxico, uma espécie de tentativa de descrição do léxico de uma língua”. A *novidade* (nem tão nova assim) nesse campo é a noção de dicionário mental.

Não conseguimos rastrear quando essa idéia surge no debate em linguística, mas podemos delimitar um marco a partir do qual investidas sobre o léxico em termos de um dicionário mental tomam corpo: o artigo de Chomsky (1970) *Notas sobre a derivação*, a partir do qual o léxico começa a ganhar *status* gerativo. Nesse artigo, lemos o seguinte:

Proponho que uma gramática possui uma base, consistindo de um componente categorial (que assumo ser uma gramática livre de qualquer contexto) e de um léxico. O léxico consiste de entradas lexicais, cada uma delas com um sistema de traços específicos.

Notemos três pontos específicos, e que nos interessam de perto:

- uma gramática livre de qualquer contexto;
- o léxico fazendo parte da base da gramática; e,
- o léxico consistindo em entradas lexicais, com um sistema de traços específicos.

Quando o autor assume uma gramática livre de qualquer contexto, isso diz respeito especificamente ao componente inato, subjacente às estruturas mentais dos falantes, que herdamos da evolução da espécie humana a capacidade, a que o autor chama de competência, para a língua numa sociedade.

Se o léxico faz parte da base da gramática, podemos, por dedução, assumir também que o léxico é de base mentalista. Não é que ele seja inato, como a gramática, mas ele se torna *mental* a partir das experiências dos iniciantes da fala numa dada sociedade.

E mais, tocados por dados de uma língua – o seu léxico –, os falantes constroem o seu dicionário a partir de critérios subjetivos. Nesses termos, cada pessoa possui o seu próprio dicionário, na medida em que se apropriaria das entradas a partir de critérios individuais, e organizaria essas entradas lexicais a partir de interesses semânticos em potencial.

E podemos estar bem certos de que cada pessoa encontra em seu léxico mental as entradas lexicais necessárias para a composição de sintagmas. Encontrar essas entradas lexicais significa jogar com uma gramática lexical sofisticada.

Se o léxico consiste de um sistema de traços específicos, poderemos nos arriscar a assumir algo temerário, mas lógico: há traços comuns para cada entrada lexical, mas cada usuário de uma língua constrói traços para cada entrada não necessariamente coincidentes com os traços construídos por um outro usuário.

Para a efetiva aceitação semântica de um item, veremos abaixo que é necessária a conjuração entre modulação de sentido e escolhas lexicais e contextuais. E esse léxico é uma força viva, atuante de maneira similar à gramática: é “(...) uma gramática de *regras vocabulares* (uma gramática de regras sintagmáticas para palavras) que possa encaixar uma palavra”<sup>13</sup> na outra ou um item num outro. Como nos inusitados dos encaixes abaixo:

**(21) -*Esse menino tem todo o sintoma do povo da política. É invencioneiro e linguarudo.***  
**(pág. 30)**

Os itens [invenc + ion + eiro] e [lingua + (r) udo] possuem “uma anatomia delicada composta de partes, chamadas morfemas que se associam de determinadas maneiras. O sistema de regras vocabulares é uma extensão do sistema de regras sintagmáticas” (PINKER, 1994, p.157).

Como funcionaria um dicionário mental? Quais propriedades podem fazer parte desse dicionário? Para cercar essas questões, vamos dialogar com Pinker (1994) e Miller (1982.), pois a delimitação de unidades chamadas lexemas necessita de uma investigação das características e da definição do léxico mental para o tipo de argumentação traçada aqui, e objetivo desta seção.

Conhecer um vocabulário é um tipo de precondição para um falante pronunciar e entender sentenças de uma língua. Diríamos que é a precondição mais imediatamente necessária para a comunicação por meio de uma língua natural. O que, então, uma pessoa aprende quando assimila um vocabulário, amplia-o, restaura-o, movimenta-se mentalmente com ele? Vamos supor que o conhecimento lexical pode ser tocado pela forma como os dicionários registram as palavras de uma língua. Os exemplos abaixo vão nos ajudar a nos lembrar de como, basicamente, esse registro é feito:

---

<sup>13</sup> Pinker, (*op.cit.*), p.156.

(22) *Esmeraldina Nogueira – Paixão de rebite de Ponciano. Dona de um par de covinhas de muita graça e olhos verdes. Branca, de cabelo em formato de labareda. Mulher de Pernambuco Nogueira.*(CL, p. 306 )

(23) **Enchaînement:** [ãsenmã] nm (**a**) (série) (d'événements, de circonstances) series, chain (**b**) (liaison) (d'idées, de séquences) linking (**c**) (em dense) sequence (of steps); (em ballet) enchaînement (**d**) Ordinat concatenation. (WEBSTER, Inglês/Francês, p.179.)

(24) **Desvaliação:** [De desvaliar + cão]. S.f. Ação ou efeito de desvaliar-(se). (Dicionário AURÉLIO, p.665)

Os verbetes que aparecem em [22], [23] e [24] foram retirados de três tipos comuns de dicionário. O [22] foi retirado de CL; o [23] de um dicionário bilíngue inglês-francês/francês-inglês; e o [24] foi retirado de um dicionário monolíngue, o dicionário Aurélio da língua portuguesa falada em território brasileiro.

Esses verbetes, no nosso caso, entradas lexicais, não são simplesmente uma sequência de itens que remetem a outros nem sentidos compartilhados de maneira simétrica, alfabética e/ou cronológica: são verdadeiros textos-colônia.

As propriedades de um texto-colônia, apresentadas abaixo por Dionísio Silva (1996, p.126) são semelhantes às propriedades para o léxico mental:

- 1.o significado não deriva de uma seqüência;
- 2.as unidades adjacentes não formam uma prosa contínua;
- 3.não há um frame contextual;
- 4.não há um autor individual identificável;
- 5.um componente pode ser usado sem referência aos demais;
- 6.os componentes podem ser reimpressos ou reutilizados em trabalhos subseqüentes;
- 7.os componentes podem ser usados, removidos, alterados;
- 8.muitos componentes podem servir a mesma função;
- 9.há uma seqüência alfabética, numérica ou temporal.

Os dicionários também indicam alguns tipos de informações associadas à palavra: pronúncia, etimologia, categoria sintática, e itens afixais fazendo parte de sua estrutura; eles ainda exemplificam os usos por meio de frases em que esses itens aparecem (sugerem contextos possíveis); comentários a respeito de seu uso e sinônimos próximos, quadros e tabelas apropriados para determinados contextos.

Se fôssemos procurar pelo lugar do léxico em teorias discursivas, encontraríamos provavelmente sugestões como as que seguem:

A memória, por sua vez, tem suas características como pensada em relação ao discurso. E, nessa perspectiva, ela é tratada como interdiscurso. Este é definido como aquilo que fala antes, em outro lugar, independentemente. Ou seja, é o que chamamos memória discursiva: o saber discursivo que torna possível todo dizer e que retorna sob a forma do pré-construído, o já-dito, que está na base do dizível, sustentando cada tomada de palavra. (PULCINELI, 2003, p.31)

As informações delineadas para um texto-colônia, e a noção de memória discursiva defendida acima, entrelaçando algo como um pré-construído, nos ajudam a entender o conceito de léxico mental.

A capacidade de manipulação do vocabulário em alguma língua exige de um falante que ele aprenda informações sobre todo e cada um dos itens que ele utiliza. Para Miller (1982), então, é conveniente supor que esse conhecimento dos itens lexicais só é possível pois esses itens compõem o que ele chama de léxico mental.

Para explorar como tais informações diversas, delimitadas aproximadamente pela citação de Dionísio e de Pulcineli acima, são integradas e representadas na memória dos usuários de uma língua, é necessário considerar qual é e como é feito esse uso pelos usuários.

Quem permite isso é o léxico mental,

(...) necessário mais imediatamente para produzir e entender sentenças, assim, a informação lexical deve ser representada de uma forma compatível com o que as pessoas conhecem sobre sintaxe e sobre semântica das sentenças em sua língua. (MILLER, *op. cit.*, p.61).

Talvez a mais importante parte do léxico mental seja o estoque passivo de entradas, disponível para uso pelos componentes mais ativos dos sistemas processuais de informação que envolvem uma língua.

Embora novas entradas possam ser adquiridas, e as palavras raramente usadas possam ser esquecidas, para muitos propósitos, o léxico mental é pensado como um fixador de dados para um arquivo. Nesses termos, o estoque é passivo apenas aparentemente, pois ele é fixado por regras intercambiáveis a todo momento em que os usuários estão a postos com a sua língua.

Pensar o léxico mental como um estoque não seria algo totalmente adequado para a sofisticação dos processos mentais, pois é contraditório com o princípio de um sistema econômico e discreto: o léxico seria infinitamente amplo e sua abrangência não comportaria as intrincadas redes de cumprimento de regras.

Não temos uma organização tipo lista telefônica lexical em nossas mentes. Temos processos e padrões, demonstrados pelo estoque, antes de termos um estoque em si. Admitamos, então, que a unidade básica do léxico mental é o lexema, sobre o qual vamos investir mais acuradamente nos parágrafos que seguem. Tomemos o seguinte dado:

(25) *Um luarão de pasto subia no distanciado.* (p. 138).

O item [luarão] que aparece em (25) vai nos ajudar melhor a acompanhar a trajetória para se delimitar um lexema. Há itens muito próximos a ele, tais como [lua], [luar] e [lunar].

A manipulação contextual certamente nos indicaria tratar-se de lexemas diferentes entre si, por reunir traços diversos entre eles. Um suposto traço afetivo/avaliativo presente em [lunar] e ausente em [lua] e [luar]; assim como um traço como satélite, presente em [lua] e [lunar] e ausente para [luar] e [lunar]. Como também luz da lua presente em [luar], mais saturada em [lunar], ausente de [lua] e [lunar], por exemplo. Mas devemos concordar: esses três itens compartilham um traço semântico em comum, acessível para qualquer e para todo usuário da língua portuguesa. Enquanto cada um deles é uma unidade lexical diferente em potencial, esse traço comum seria o lexema.

Em termos morfológicos, um lexema pode ser visualizado a partir de afixos flexionais e derivacionais. Vamos acompanhar o que diz Cruse (1992, p. 77) a respeito disso:

Um afixo flexional é um elemento gramatical, pertencendo a um conjunto fechado, que pode funcionar como um componente de uma palavra: [lua + (r) + ão]; [lua-r]; [invenc + ion + eiro]; [língua + (r) + udo], (os elementos em negrito são todos afixos). Cada afixo é obrigatoriamente anexado a um esquema contendo ou consistindo de um conjunto de itens abertos (os elementos sublinhados nos exemplos acima) (...) os afixos derivacionais produzem novas unidades lexicais.

Itens de um conjunto aberto e fechado da língua serão delimitados com mais acuidade mais adiante, remetemos o leitor para lá, para maiores considerações. Estamos *visualizando* os lexemas por dentro, o observando o que eles *levam* consigo como traços identificáveis nas unidades lexicais.

Vamos retomar os verbetes [1], [6] e [7] da citação de Dionísio para entender como operam aproximadamente os lexemas: “1. o significado não deriva de uma seqüência; 6. os componentes podem ser reimpressos ou reutilizados em trabalhos subseqüentes; 7. os componentes podem ser usados, removidos, alterados...” Dionísio bem delimita para nós o jogo da construção e da estruturação de um lexema. Vejamos:

1. ‘O significado não deriva de uma seqüência’: isso significa que o usuário de uma língua não junta matematicamente dois ou mais itens para estruturar um lexema. Ele dá um *giro de 360° graus* em termos semânticos; opera ao mesmo tempo com as padrões e com o conhecimento dos itens; seleciona contextos possíveis e modula semanticamente o lexema; precisa saber antecipar e sequenciar itens; e ele faz tudo isso ao mesmo tempo.

6. ‘Os componentes podem ser reimpressos ou reutilizados em trabalhos subseqüentes’: é nesses termos que o sistema e seus itens é estável e discreto. O usuário conta com uma estabilidade semântica dos itens, assim como uma capacidade de operacionalização discreta, pois, de um conjunto finito de itens, ele consegue alto poder de formulação e reformulação de sentenças.



7. Por fim, ‘os componentes podem ser usados, removidos, alterados’: supomos uma *imprevisibilidade domada* do léxico. A expansão do léxico torna imprevisível para o usuário as ricas possibilidades de criação, e os dados do romance CL com que estamos trabalhando aqui ilustram bem isso: os itens são inesperados e mais ou menos não familiares aos usuários da língua portuguesa. Só que essa imprevisibilidade é *domada* por padrões os quais os falantes não podem deixar de cumprir. Para ver alguns deles demonstrados, remetemos o leitor a Rocha (1998).

A organização do léxico mental pode colher desta divisão – itens abertos e fechados da língua – uma maneira de ordenação desses itens. Assim como um dicionário comum não arrola em ordem alfabética, por exemplo, itens gramaticais, somente os itens lexicais, deve haver na mente um lugar diferenciado para itens lexicais e itens gramaticais. Trocando em miúdos, quando uma unidade lexical está em uso, apesar de os itens gramaticais comporem parte significativa para a sua decodificação, os falantes criam expectativas mais elaboradas em torno dos itens abertos da língua, ou seja, em torno das unidades lexicais.

Podemos, então, redefinir um lexema da seguinte forma: um lexema é um conjunto abstrato de traços semânticos, gramaticais e contextuais que podem residir numa dada entrada lexical antes de se transformar em unidade lexical. “O lexema antecede a unidade lexical”. Haveria um conjunto de unidades lexicais para um único lexema. O esquema abaixo pode nos ajudar no que queremos dizer quando afirmamos “conjunto abstrato de traços semânticos, gramaticais e contextuais”. Entendendo aqui que desses traços também fazem parte informações gramaticais. Vamos retomar a sentença que aparece em (19):

(25) *Um luarão de pasto subia no distanciado.* (p. 138).

E pensá-la nos seguintes termos, como na sequência em 20:

(26) Um \_\_\_\_\_ de pasto subia no distanciado.

O espaço ( \_\_\_\_\_ ) em (26) exige algumas especificidades em termos de traços semânticos e gramaticais/categoriais para ser adequadamente preenchido na atualização dessa sentença em termos funcionais. Os falantes não recorrem a um depósito onde estão armazenados os itens na sua memória, simplesmente. Ao recorrer a uma memória e reconstituir um pré-construído, ele leva algumas senhas para ter acesso a esse depósito:

- a) informações gramaticais/contextuais: em (25), um item no singular e no masculino; um item que controle número e gênero do modificador que o antecede, o artigo *um*; um item que funcione como núcleo do sintagma;

- b) informações semânticas: algo como [lua]; e
- c) informações contextuais: toda a sentença construída em (25).

Ao atualizar no espaço em branco o item como ele aparece em (25), o falante faz a transposição das abstrações em termos de lexema para sua função em termos de item lexical.

Um critério alternativo para assinalar as unidades lexicais em sua contraparte como lexema é que os sentidos podem pertencer a um espectro semântico, nas palavras de Cruse (1992). Vamos reproduzir aqui a sentença (17) que aparece logo acima, e explicitar esse espectro semântico a que se reporta esse autor:

**(17) - *Aquilo é tábua de passar roupa. Moça para o primo tem que ter coxão fornido, capaz de agüentar os repuxos. (CL, p. 17)***

Mais uma vez, o item em jogo é [*tábua de passar roupa*]. O contexto imediatamente posterior a esse item seguramente nos indica que se trata de um qualificador avaliando a moça citada no excerto. Há um espectro que pertence a [*tábua de passar roupa*] que qualifica moça, avaliando sua magreza. A moça não é o objeto onde se passam as roupas – não daria para passar roupa em cima da moça. A moça rouba um traço (ou mais de um desse objeto), sua espessura, por exemplo, e a ela é comparada, seria algo em torno de: “A moça é tão magra e dura como uma tábua de passar roupa”, a que acrescentaríamos o dado avaliativo/depreciativo dessa moça em questão. Vamos manipular uma outra sentença (27) onde apareça este item [*tábua de passar roupa*].

**(27) *A moça era tão difícil de se ver quanto aquela tábua de passar roupa escondida no sótão.***

Em (27), em jogo temos uma cena em que moça também está sendo qualificada e/ou comparada a um item [*tábua de passar roupa*]. Veja, porém, que o traço que moça “rouba” de [*tábua de passar roupa*] já não é mais a sua espessura, mas o fato de esse item estar difícil, escondido.

Há um espectro semântico encontrado no item [*tábua de passar roupa*] tanto em (17) quanto em (27), mas é como se esse espectro fosse prismático: arrolado de ângulos tão diferentes que tornaria esse item comum apenas em termos reduzidos: reduzidos à sua condição de lexema.

Dentre as características de um lexema, haveria aquelas que seriam mais básicas ou centrais do que outras, que seriam mais periféricas. Vamos tentar descobrir quais seriam elas, e classificá-las, nos parágrafos que seguem.

Há características mais estabelecidas e algumas outras menos estabelecidas compondo o escopo de um lexema.

Seria ideal esperar de um dicionário que elencasse as características essenciais de um lexema, e se essas características permanecem idênticas quando autores de diferentes dicionários as definem como as mesmas. Podemos, então, comparar ao que se passa em termos do léxico mental e os lexemas que habitam nele.

Tomemos dois verbetes similares que aparecem nos dicionários Aurélio e Houaiss, e esperamos que esses dois lexicógrafos representem dois falantes hipotéticos, e cada um deles com o seu próprio léxico mental. Dessa forma, queremos reconstruir um léxico mental a partir de sua comparação com um léxico arrolado nos dicionários de uso de língua disponíveis no mercado:

(28) *Legista1: S.2.g.V. Médico-legista.*

*Legista2: [Do lat. med. legista]. Adj. 2 g. 1. Que conhece ou estuda as leis. S.2 g. 2: pessoa que conhece a fundo as leis; legisperito. (AURÉLIO, p.1191).*

(29) *Legista1: adj.2g.s.2g. (sXV cf.FichIVPM) JUR especialista em leis; ETIM fr. légiste e, este, do lat.medv.legista, der. de lex, legis, lei; ver legi(s)-; f.hist.sXV legistas, sXV llegendas.*

*Legista2: s.2g. JUR MED red. de MÉDICO-LEGISTA. (HOUAISS, p. 1736)*

É possível delinear algo comum:

- a) As informações gramaticais permanecem as mesmas, e seria impossível que elas assim não permanecessem já que se trata de um item da língua portuguesa.

Conclusão provisória 3: em termos de domínio de regras gramaticais podemos afirmar que nossos dois falantes hipotéticos dominam as mesmas regras; seria impossível que eles não as dominassem, pois, além de estarem se comunicando por meio de uma língua, eles se comunicariam por meio de um código similar, a língua portuguesa.

- b) Se eles apontam itens idênticos na reconstrução de sentido, é sinal de que esse item possui um escopo comum recobrando seus sentidos; para o nosso falante hipotético, é aquilo que as teorias do discurso chamam de pré-construído ou memória discursiva compartilhada. Isso aponta também para algo delineado em parágrafos anteriores: a estabilidade semântica desse item, ou sua relativa estabilidade semântica, dado que até agora o nosso objetivo é estritamente semântico, mas poderíamos recorrer à ideologia, que nos ajudaria a elasticar o escopo delimitado tanto para os lexicógrafos quanto para os nossos dois falantes hipotéticos;
- c) Se eles identificassem (o que não é o caso) contextos similares para a ocorrência desse item, é sinal de que o dado funcional também joga com peculiaridades para defini-lo; e os

nossos falantes hipotéticos, por compartilharem uma mesma comunidade de fala, estariam aptos a reconstruir tantos contextos possíveis para esse item.

Se quisermos hierarquizar as propriedades semânticas de um lexema, poderíamos compará-lo a uma unidade lexical nestes termos: há *unidades lexicais primárias* e *unidades lexicais secundárias* de um lexema.

Não queremos hierarquizar em termos de uma dicotomia, as primárias anulando e/ou substituindo as secundárias, ou uma sendo mais importante de que a outra. Em verdade, elas atuam em conjunto para algumas tarefas, explicitadas abaixo, como também perfazem um *continuum* semântico. Faremos considerações em relação a elas nos parágrafos que seguem.

Todos os falantes reconstituem contextos possíveis quando está em jogo o contrato semântico. Eles cooperam entre si para levantar as reais necessidades semânticas de uma sentença e, assim, comunicar-se plenamente. Eles realizam algumas tarefas previsíveis para a significação efetivar-se com sucesso:

- fazem esforços para desambiguar mentalmente as possíveis ambiguidades contextuais e lexicais;
- empreendem uma busca em suas memórias para saber onde e como atualizar semanticamente os itens;
- limpam o *terreno semântico* para acomodar a interpretação consequente e necessária para uma sentença funcionar plenamente ou parcialmente e, dessa forma, efetiva-se a comunicação por meio de uma língua natural;
- atuam num campo semântico, de onde conseguem retirar dele o necessário para recompor os sentidos.

Se de um lexema podemos inferir unidades lexicais primárias, os exemplos abaixo podem nos ajudar a digerir melhor esse conceito:

(30) *Morreu Fonseca, morreu o ano. Abri mão de várias regalias, menos do charuto.* (CL, p. 282)

Uma unidade lexical primária pode ser pensada em termos de relações semânticas ou metonímicas possíveis para qualquer contexto e/ou independente de qualquer contexto, do tipo: [*Fonseca: nome próprio*]. Independente de qualquer contexto ou para qualquer contexto, Fonseca é um nome próprio. Relutei em colocar entre esses colchetes o dado ser humano ou pessoa, pois esse nome próprio poderia servir para identificar uma rua, um cachorro, uma

cidade etc. Até mesmo em contextos altamente metafóricos é possível delinear uma interpretação de primeiro plano.

Uma unidade lexical pode ser pensada em termos contextuais, anafóricos, remissivos, sequenciais, catafóricos, do tipo: [*regalias:charuto*]. As unidades primárias, interpretadas em primeiro plano<sup>14</sup> significam exatamente aquilo a que se referem. Ninguém em sã consciência poderia negar que [*regalias*] e [*charuto*] estão de certa forma envolvidos nessa sentença, e que um remete ao outro diretamente.

Para explicitar melhor o que significa uma unidade lexical pensada em termos de segundo plano, vamos tomar de exemplo também o item [*charuto*] que aparece em [30].

Em nenhum momento naquele contexto charuto pode ser pensado como que fazendo parte de um campo de sentido do tipo [*fumar faz mal à saúde*]. É possível pensar assim, mas isso não pode ser atribuído diretamente àquele contexto.

Haveria usuários da língua que, com razão, poderiam discordar de que um charuto seja uma regalia, e que, em verdade, é prejudicial a quem fuma; mas, convenhamos, *em primeiro plano*, charuto não ocupa o campo [*fumar faz mal à saúde*] para aquele contexto.

Nesses termos, um dicionário de uma língua deve ser visto com reservas. Procuramos em dois dicionários pelo verbete [*charuto*], e o que encontramos não condiz com a interpretação de primeiro plano que aqui queremos instituir como relevante para a teoria semântica com que estamos trabalhando na descrição da noção de entrada lexical:

(31) *Charuto*: [Do *Ingl. Cheroot* < *tâmil shuruttu* 'rolo de tabaco'] *S.m.* 1. Rolo de folhas secas de fumo, preparado para fumar-se. 2. *P. ext. Cul.* Bolo outro alimento em forma de charuto. (AURÉLIO, p. 452.)

(32) *Charuto*: *s.m.* (1802 *cf. JM3*) 1 rolo alongado de folhas de tabaco para se fumar, de cor castanha, e que tem como invólucro uma folha inteiriça. 2 *p. pna.bolo* ou qualquer outro do feitio de um charuto. 3. Bebida preparada com vinho e mel. 4 *B pej.* indivíduo de cor negra 5 *AL m.q. MATA-CACHORRO* ('*servente de circo*'). (HOUAISS, p. 694.)

Como observamos das deduções acima, para uma unidade lexical de primeiro e de segundo planos, e para o que trazem os dicionários monolíngues do português, não há referência alguma em (30) que nos leve a acreditar na abrangência, e até mesmo na seriedade

---

<sup>14</sup> Definiremos primeiro plano para uma semântica no capítulo "A noção de Entrada Lexical na estrutura argumental".

semântica dos dicionários. Não custa repensar nas palavras de Biderman (op.cit), citadas logo acima: nos dicionários não está o sentido das palavras, não estão todos os sentidos possíveis, como também não há consenso para os sentidos das palavras encontrados neles.

É uma tentativa de arrolar sentidos possíveis. Ou seja, não passa de uma mera tentativa. Nesses termos, podemos afirmar que os sentidos arrolados pelos lexicógrafos como Houaiss e Aurélio em seus dicionários são os sentidos que o Aurélio e o Houaiss querem dar às palavras arroladas nos dicionários deles.

Arrolamos aqui as palavras de Biderman (*op. cit.*, p. 132), com as quais estamos de acordo, ao menos em um campo: os dicionários *flagram* um momento de uma dada cultura ou os valores culturais de uma dada sociedade, ou ainda as escolhas ideológicas, linguísticas e políticas dos lexicógrafos considerados referências para uma língua:

Um dicionário é um produto cultural destinado ao consumo do grande público. Assim sendo, é também um produto comercial, o que o faz diferente de outras obras culturais. É preciso considerar igualmente que o dicionário deve registrar a norma lingüística e lexical vigente na sociedade para a qual é elaborado, documentando a práxis lingüística dessa sociedade.

Descontados os otimismo da autora, que mais adiante compara o léxico a um tesouro de uma língua, e diz que os dicionários incluem “a nomenclatura de *todos* os conceitos lingüísticos e não-lingüísticos e de *todos* os referentes do mundo físico e do mundo cultural”, isso não é bem verdade do ponto de vista da estrutura do léxico mental.

Em termos de regras e relações até que isso é possível. Mas em termos de registro real do léxico e dos seus itens lexicais, seria improvável que num dicionário coubesse todo o léxico possível.

Estamo-nos referindo tanto ao léxico público, quanto ao léxico mental ou cognitivo, como também ao dicionário público e ao dicionário cognitivo ou mental. Na verdade, cada um de nós apenas mantém um conjunto de itens lexicais possíveis e de interesses que estão próximos à nossa comunidade seja de fala, científica, religiosa, ideológica etc.

Não concordamos também com a autora que afirma ser o léxico o “menos lingüístico de todos os domínios da linguagem”. Acreditamos justamente no contrário: o léxico é “o” componente lingüístico por excelência, por três razões:

- sem o léxico não haveria a língua;
- o léxico pode ser estruturado da mesma forma que a sintaxe traduz a língua, em termos de regras e concatenações, hierarquias e motivações; e,
- o léxico mantém relações semânticas e sintáticas da mesma forma que outros domínios da gramática.

## 4.5 O PRINCÍPIO DAS RELAÇÕES RECORRENTES

Vamos introduzir agora um novo conceito para compor o objetivo a que estamos nos propondo neste capítulo: estudar os lexemas. O *princípio das relações recorrentes*. Essas relações de recorrência são simultaneamente gramaticais e semânticas. Nós nos reportamos a elas rapidamente no começo deste capítulo, mas agora vamos investir um pouco mais de esforço em caracterizá-las.

Vai nos ajudar o que ficou conhecido na literatura linguística como semântica dos papéis temáticos. (CANÇADO, 2003; FRANCHI, 1996; MIOTO et alli, 2002; JAKENDORF, 1978)

Não é nosso propósito discutir exaustivamente o que são esses papéis temáticos; apresentar os papéis temáticos para a língua portuguesa; problematizá-los no sentido de apontar suas falhas ou seus limites.

Queremos apenas relacioná-los dentro do quadro teórico da semântica lexical, ao mesmo tempo em que retomá-los para nos servir como suporte de explicitação de um princípio de relações semânticas recorrentes. Vamos começar apresentando uma definição encontrada num trabalho de Cançado (1995, p. 83):

A Teoria dos Papéis Temáticos construiu-se inicialmente para expressar a estrutura argumental dos verbos, sob a hipótese de que as distintas funções semânticas associadas aos argumentos do predicado verbal seriam relevantes para determinar a estrutura sintática da oração e não funções sintáticas como as de sujeito e objeto.

Esse excerto de Cançado leva-nos a refletir sobre as funções semânticas de um item (a autora se refere especificamente aos verbos), e suas forças simbólicas na determinação da estrutura de uma sentença. Vejamos agora o que encontramos em Miotto et alii (1998, p.119), discorrendo sobre o mesmo assunto:

Todos sabemos que para além de dominarmos as regras de formação sentencial, é impossível falar uma língua sem dominar o seu léxico, isto é, o conjunto de palavras que constitui o dicionário da língua em questão. Mais do que isso: as palavras da língua têm propriedades tais que o aparecimento de um certo item lexical já nos faz esperar um outro item ou grupo de itens.

Ao avaliarmos a necessidade de pensar o léxico como organizado em um dicionário mental, vimos que as informações semânticas contidas nos itens lexicais, e disponíveis nesse dicionário, são muito mais que relevantes tanto para a construção de sentenças como para a formulação de contextos.

Retomemos agora o princípio das relações recorrentes. Assumimos que, para explicitar melhor esse princípio, informações semânticas e gramaticais atuam em conjunto. Pois bem.

Está inscrito num item lexical, como se estivesse gravado em seu DNA lexical informações de sentido como também informações categoriais. Vamos observar como opera esse princípio no conjunto de sentenças que seguem:

(33) *Recriminei o covardismo dele...* (CL, p. 47)

(34) *O covardismo dele recriminado.*

(35) *Aprendi o caminho* (p.91).

(36) *Caminho aprendido.*

(37) *Inventei compromisso de mulata teúda e manteúda de propósito.* (p.109)

(38) *Compromisso de mulata teúda e manteúda inventado de propósito.*

Há um evidente paralelismo sintático entre os itens: [*Recriminar* : *covardismo*] em (33) e (34); [*Aprender* : *caminho*] em (35) e (36); e [*Inventar* : *compromisso*] em (37) e (38).

Esse paralelismo sintático joga com posições diferentes para esses itens nas sentenças que vão de (33) a (38): mas há um princípio de relações recorrentes entre os itens em pares assinalados aqui. Há relacionamentos semânticos que se mantêm, dadas as informações semânticas contidas nos lexemas e as informações gramaticais inscritas no DNA lexical de cada um desses itens.

Os lexemas são famílias de unidades lexicais. Uma unidade lexical é a união de um sentido isolado com uma forma lexical; uma forma lexical é uma abstração deduzida por um conjunto mais amplo, e o princípio das relações recorrentes de sentido pode relacionar unidades lexicais e lexemas.

Considerar o léxico como um componente relevante, com o mesmo *status* para descrição e análise das línguas que a gramática e a sintaxe já desfrutam há muito tempo, parece hoje algo quase natural para uma teoria linguística.

Nenhuma descrição consequente e necessariamente aprofundada das línguas naturais pode deixar o componente lexical de lado, sob pena de suas descrições e análises tornarem-se vagas ou improcedentes. Mas essa descoberta não tem tanto tempo assim. Podemos até datá-la. O artigo de Noam Chomsky, “Remarks on nominalization”, já citado aqui, abriu possibilidades investigativas nunca antes imaginadas para o léxico, e trouxe consequências fundamentais para estabelecer uma teoria consistente em semântica lexical.

O conhecimento lexical entre os falantes de uma língua natural assemelha-se ao conhecimento gramatical. Assim como falar uma língua é dominar regras e jogar com elas, o conhecimento do léxico procede em bases similares: conhecer o léxico de uma língua é conhecer regras e também jogar com elas. As informações veiculadas por, e inerentes a, uma



sentença de uma língua devem sua estrutura e organização, em grande parte, ao conhecimento lexical que os falantes possuem de sua língua.

A decodificação semântica de uma sentença começa com uma perspectiva que os falantes possuem de um conhecimento do léxico. Em linguística, a semântica lexical é um campo teórico que se ocupa basicamente da relevância do léxico para uma língua, e os conhecimentos teóricos investigados nesse campo podem englobar os seguintes processos:

- Explicitar a natureza multifacetada, polimórfica, prismática e polissêmica da língua.
- Caracterizar a semanticidade das sentenças de uma língua natural e o envolvimento de fatores cognitivos demarcadores dessa semanticidade.
- Capturar a criatividade lexical dos falantes de uma língua em contextos a cada momento novos e diferentes.
- Desenvolver uma representação do léxico em termos relacionais.

Esses processos são, basicamente, os objetivos de uma semântica lexical consistente. Antes de prosseguir o caminho para cercar esses objetivos, gostaríamos de estabelecer dois pontos fundamentais para nossas discussões:

- a) Sem levar em consideração a estrutura sintática de uma língua, um estudo em semântica lexical pode ser fadado ao fracasso. Divorciar o estudo dos sentidos da estrutura da língua, além de ser impossível, é impraticável, pois a estrutura das sentenças direciona os caminhos de interpretação, e vice-versa. As distinções gramaticais são geralmente muito úteis na medida em que direcionam para aquilo que compete a uma teoria semântica do léxico.
- b) Os sentidos das palavras refletem de alguma forma as estruturas conceituais mais profundas do sistema cognitivo e o domínio em que esse sistema opera. Isso equivale a dizer que a semântica de uma língua natural seria a imagem de princípios organizacionais conceituais não-linguísticos, qualquer que seja a estrutura da língua. Em outros termos, uma língua em uso se vale de palavras, de algo que não seja palavras, de lapsos, de gestos, de vazios, elipses etc. A atuação do léxico também deve muito à natureza do sistema cognitivo, mas também deve sua efetividade a alguns princípios.

Vamos enumerar esses princípios no seguintes tópicos:

Tópico 1: *Para caracterizar semanticamente um item em termos da semântica lexical, precisamos observar o que há de não-semântico nesse item, em relação a esse item ou em relação a um contexto sentencial do qual esse item emerge.*

Vamos pressupor que o léxico se assemelhe a um dicionário, e que todos nós temos um dicionário, o nosso dicionário particular, em nossas cabeças. Para formar sentenças bem formadas semanticamente vamos a esse dicionário e procuramos em nossos arquivos as entradas lexicais necessárias ou adequadas ou possíveis. Estamos no campo da caracterização de sentido de uma entrada e utilizando as informações cognitivas para revirar esse arquivo, retirar dele as entradas necessárias e descartar outras. Precisamos juntar essas entradas para formar as sentenças.

Veja que nós estamos utilizando duas frentes para formar uma sentença: estamos conjugando esforços cognitivos e semânticos. Por esforços cognitivos vamos pressupor informações e noções abstratas longe de influências semânticas. Por exemplo, podemos sugerir que fatores discursivos e pragmáticos manipulariam diferentemente e/ou separadamente e/ou simultaneamente as contribuições semânticas dos itens lexicais.

Embora isso não seja algo categórico – contextos não semânticos (em estado de dicionário) só podem ser explicitados semanticamente – isso vai nos ajudar a focalizar algo imprescindível numa teoria em semântica lexical: regras e relações entre itens numa sentença dada.

O que faríamos com um item retirado do arquivo-dicionário-mental para construir uma sentença, se não soubéssemos relacionar e/ou manipular semanticamente esse item em termos de regras de manipulação do léxico? O que faríamos se não conseguíssemos encaixar as peças dos itens uma a uma e culminar numa sentença bem formada? Embora isso fuja ao nosso escopo, não pode ser abandonado como o contraponto para uma descrição em semântica lexical que se pretende consequente.

Tópico 2: *Embora o dado não-semântico e a gramática sejam imprescindíveis para a fixação dos sentidos possíveis para uma entrada lexical, queremos demarcar que existe algo para além dessas assertivas o qual delimitaremos como fronteiras demarcadoras de uma semântica lexical.*

Queremos argumentar que é possível estabelecer um método de decomposição lexical próprio a uma semântica lexical. Isso pressupõe que o nosso trabalho aqui pode alcançar:

- a) uma teoria recursiva que atua no léxico – assim como o princípio da recursividade atuaria para a gramática –: os itens não morrem, renascem em outros itens; os itens

podem sofrer bloqueios na elaboração de outros itens mas também podem romper esses bloqueios;

- b) algo como uma boa formação semântica para as sentenças; e,
- c) recursos para variados níveis de interpretação semântica, que desenvolveremos ao estabelecer as diferenças entre lexemas e itens lexicais.

*Tópico 3: Todas as categorias sintáticas delimitadas para uma sentença estão relacionadas a uma motivação semântica.*

Isso significa dizer que no léxico encontramos informações subjacentes – os traços semânticos – que distribuiriam as palavras em classes gramaticais, categorias sintáticas, predicadores, núcleos de sintagmas etc. Remetemos o leitor para nosso trabalho “Gênero linguístico: semântica e função” (2007), em que pressupomos a categoria gênero linguístico como essencial para a delimitação de um constituinte semântico em português.

Os trabalhos para os verbos são bem desenvolvidos. A noção de argumentos de um verbo e suas relações semânticas estão bem estabelecidas, mas argumentamos que **todas** as categorias são passíveis de uma checagem e de uma demanda semântica, dado que as relações semânticas são inerentes a elas.

A gramática tradicional sempre reservou um lugar para a semântica, tão destacado quanto aquele da morfologia e da sintaxe.

Aprendemos que a semântica é o campo interessado em estudar, fixar, compreender algo, muitas vezes vago e impreciso, chamado sentido, mais especificamente, o sentido das palavras, como vimos acima. O sentido é percebido como uma evidência desde o processo de aquisição de uma língua natural em terna idade.

Mas conceber o sentido como algo natural no cotidiano de um falante e como objeto de pesquisa no campo linguístico gera controvérsias e rupturas até hoje inconciliáveis para semanticistas e gramáticos.

Para os gramáticos, por exemplo, as relações de sinonímia e antonímia são a base de seu capítulo Semântica em uma gramática tradicional. Homonímia e polissemia também são conceitos exaustivamente trabalhados nesse capítulo. Nos livros didáticos do ensino fundamental, é comum acompanhar um texto, um vocabulário com as possíveis palavras difíceis para os alunos, e seus respectivos significados. Mas há uma distância considerável entre o que os falantes pressupõem como significado, o que preconizam as gramáticas

tradicionais sobre esse tema e o que a pesquisa em linguística delimita como os interesses do campo semântico.

A sinonímia, por exemplo, é um mistério ainda não desvendado. Será que existiriam sinônimos entre os itens lexicais? Em outras palavras será que caberia num mesmo contexto dois itens como [*menino* e *garoto*]? Sendo [*menino*] sinônimo de [*garoto*]? Vamos sugerir as seguintes sentenças:

(39) *O menino pegou a bola.*

(40) *O garoto pegou a bola.*

Trata-se de um mesmo contexto, dele fazendo parte dois itens sinônimos [*menino* e *garoto*], ou (39) e (40) são contextos diferentes, pois preenchidos por itens, no mínimo, morfológica e fonologicamente diferentes? Alguns radicalizam e dizem que sinônimos não existem na língua. O mesmo pode ser dito para as relações de antonímia.

Vamos acompanhar as sentenças abaixo:

(41) – *Você vai ao cinema comigo Joana?*

– *Sim, já estou indo!* (dito num tom irônico e de desprezo para com quem fez a pergunta).

(42) – *Você vai ao cinema comigo hoje, Joana?*

– *Não, não vou!*

Será que [*sim*] e [*não*] são antônimos entre si, assumindo que a relação de antonímia é uma relação de contrários e oposições? Se assumirmos que seja isso, será que [*sim*] é o oposto ou o contrário de [*não*] nos contextos (42) e (43)?

Ao nos depararmos com questões simples como essas, podemos, então, deduzir que uma investigação científica em semântica tem tido sérios problemas a enfrentar. E o trabalho pioneiro de Michel Bréal (1883), ao assumir o sentido como um elemento especificamente linguístico, já entrou para a história.

Amplamente resenhado nos estudos semânticos, esse trabalho de Bréal é o precursor dos estudos em semântica lexical como nós os conhecemos hoje. Quando dizemos precursor, estamos fixando que Bréal apenas toca, de leve, em algo caro à semântica lexical: o sentido como algo autônomo para uma teoria linguística.

Vamos rapidamente nos reportar a alguns pressupostos encontrados nesse trabalho: “O homem nada influencia no desenvolvimento da linguagem e que as palavras – forma e sentido – levam uma existência que lhes é própria”. Essas palavras desbravadoras dos estudos semânticos modernos fixam alguns pressupostos:

- os sentidos das palavras estão fora de nós;

- não temos capacidade de significar, pois já encontramos os significados prontos e acabados.

Passados mais de cem anos desde o momento em que essas palavras foram ditas, ainda não conseguimos encontrar uma maneira de refutá-las por completo do ponto de vista da investigação linguística em semântica.

Retomamos noções sobre semântica com o propósito de fixarmos conceituações sobre a noção de lexema. Oscilamos entre três caminhos em termos de estabelecimento de uma base teórica em semântica:

- os sentidos são arbitrários e independem da ação das mentes humanas sobre eles;
- as palavras não têm sentido em si; sendo os seus sentidos estabelecidos no frescor da constituição contextual das frases e textos que produzimos; e
- as palavras gozam de uma estabilidade precária de sentidos.

Mas há fronteiras de vários tipos entre esses três pressupostos e suas relações com a concepção de lexema. Um lexema, basicamente, é coberto por uma gama de noções deste tipo:

- a) fronteiras relacionais entre os itens numa sentença;
- b) distribuição de itens em um dicionário – o léxico;
- c) comparação dos itens entre si;
- d) regras de encadeamento de itens;
- e) traços contidos em um item;
- f) informações tipo DNA lexical para cada item;
- g) visualização de itens como entradas;
- h) investidas gramaticais sobre um item.

O que enumeramos de (a) a (h) faz parte dos domínios e fronteiras de uma teoria semântica. Vamos desdobrá-los rapidamente aqui:

- a) os itens [carro] e [correr] se relacionam de uma maneira em uma sentença, enquanto os itens [de] e [casa] se relacionam de uma outra forma:

(41) *O carro corre perto de casa;*

- b) as sentenças são formadas de itens retirados de um léxico, e concatenadas segundo regras de organização delimitadas pela gramática da língua;
- c) os itens [homem] e [mulher] compartilham significados mais próximos do que os itens [rocha] e [pedra], e o léxico os distribuiria de uma maneira diversificada;

- d) os itens [guarda] e [chuva] cumprem regras de formação de palavras que possibilitariam a criação de um item complexo como [guarda-chuva], e vetariam a criação de um item do tipo [\*paredor], uma suposta junção de itens como [parede] + [dor];
- e) os itens [choveu] e [chuva] guardam diferenças em termos gramaticais pois incidem sobre eles traços gramaticais diferentes: o traço +verbo para [choveu] e o traço + nome para [chuva];
- f) há informações do tipo X no item [garota] – gênero e número, por exemplo, diferentes daquelas do tipo Y no item [garotadas] – por exemplo, derivação e número plural;
- g) os lexemas podem ser organizados no léxico como entradas, por exemplo: o item [garoto] teria uma entrada do tipo diferente de [banco] num dicionário virtual; e
- h) há investidas gramaticais em um item do ponto de vista paradigmático – o item visualizado isoladamente; e sintagmático – o item delimitado numa estrutura de uma sentença.

Um outro problema para os domínios de uma semântica são conceitos como palavra, frase, texto e discurso. Esses conceitos estão inter-relacionados de tal forma a um item lexical, que teríamos o significado de item lexical diferente quando lidamos com textos, parágrafos e discursos, ou poderíamos assumir um mesmo conceito para eles?

Num trabalho famoso de Kempson (1978, p. 17), buscamos algumas bases para a delimitação exaustiva de uma teoria semântica, e, em particular, o lugar do léxico dentro dessa teoria. A autora assume que:

(...) uma teoria semântica deve satisfazer pelo menos três condições: (i) deve apreender, para qualquer linguagem, a natureza do significado de palavras e de sentenças, e explicar a natureza da relação entre eles; (ii) deve ser capaz de prever as ambigüidades nas formas de uma linguagem, sejam em palavras ou em sentenças; (iii) deve caracterizar e explicar as relações sistemáticas entre palavras e entre sentenças de uma língua – isto é, deve proporcionar uma explicação explícita das relações de sinonímia, inclusão lógica, acarretamento, contradição etc. Qualquer teoria que deixe de captar essas relações, seja geral ou nos casos particulares que façam previsões errôneas, deve ser inadequada em princípio ou em alguns de seus detalhes teóricos.

Quando a autora propõe “apreender a natureza do significado de palavras e de sentenças, e explicar a natureza da relação entre eles” isso significa uma investida no léxico da língua e seus desdobramentos para uma teoria semântica que satisfaça uma condição de exaustividade e abrangência.

O problema é como captar essas relações como também explicitar a sua natureza, e é isso o que vamos tentar a partir dos parágrafos que seguem, baseados no estudos de cognição semântica de Miller (*op.cit.*).

Para esse autor, o léxico é o ponto de partida de qualquer estudo que se queira exaustivo em semântica, para isso, deixa bem claro que uma semântica lexical não se funda sem relacionar-se diretamente com uma base sintática, e propõe que:

O objetivo principal de qualquer teoria semântica lexical é classificar adequadamente os itens lexicais de uma língua nas classes predicativas de sua expressão sintática e semântica. Além disso, essa teoria não deve meramente mapear os sentidos dos itens lexicais para uma sentença, numa base individual. Antes, deve capturar as relações semânticas entre as palavras, de uma maneira tal que facilite esse mapeamento.<sup>15</sup> (MILLER, *op. cit.*, p.8)

Pelas palavras de Miller, podemos deduzir que uma teoria semântica lexical jamais pode ser meramente descritiva. Antes de mais nada, compreende uma capacidade de análise e investigação que dê conta de forma concatenada e exaustiva das relações entre semântica e sintaxe de uma maneira diferente daquela tarefa legada a uma teoria dos papéis temáticos, a qual aludimos logo acima.

Vamos, então, continuar com esse autor, observando como esse mapeamento pode ser feito, ou então, acatar a sugestão alternativa de Kempson (*op.cit.*) e observar como podemos apreender a natureza do significado das entradas.

Podemos propor alguns esquemas para captar a natureza dos significados das entradas:

- a) a designação *in situ*, direta e inseparável de relações vividas e chamadas de designação vaga (...): substantivo designativo (próprio ou comum, mas sem artigo), outra designação empregada designativamente;
- b) a *lexicalização*, que permite separar as denominações de seu uso designativo e memorizá-las como signos lexicais, integrados aos sistemas de relações que definem opositivamente os seus significados;
- c) a *gramaticalização*: dissociam-se, por meio de marcas formais específicas termos e operadores ou relatores. Desse modo, se estabelece um sistema combinatório abstrato, dependente apenas de suas regras de organização interna e inclinado a uma generalização ideal.

---

<sup>15</sup> It is the goal of any lexical semantic theory to adequately classify the lexical items of a language into classes predictive of their syntactic and semantic expression. Furthermore, such a theory should not merely map the meanings of lexical items per sentence, on an individual basis. Rather, it should capture the semantic relations between words in such a way which facilitates this mapping. (Tradução nossa).

A natureza dos significados das entradas lexicais é sempre a busca por considerar padrões de lexicalização constantes entre itens e suas relações possíveis na memória dos falantes.

A natureza da gramaticalização é um outro problema que foge ao escopo de nosso trabalho.

O que nos interessa agora é a natureza do significado sob sua designação *in situ*, e suas relações diretas e inseparáveis do vivido. Quem primeiro toca nessas questões em semântica é o campo lógico da semântica formal. Próximas considerações a ser feitas.

A forma com que o léxico se expande e se estabelece numa língua influencia o desenho e as regras da gramática de uma língua, como veremos logo abaixo, através do preenchimento de argumentos de um verbo proposto pela teoria dos papéis temáticos. Mas, o ponto central da semântica lexical, queremos crer, tem sido uma busca por delimitar seus princípios em termos lógicos e mapear suas representações sintáticas em termos sentenciais em direção a sua representação lógica.

Talvez o campo de estudo semântico pioneiro de análise tenha sido a semântica formal. Dentro desse campo, o aspecto fundamental do sentido das palavras talvez seja o seu tipo semântico, delimitado como informação categorial. A informação categorial ou o tipo de informação categorial determina não somente como uma palavra se comporta sintaticamente, como também a quais elementos a categoria se refere semanticamente. Por exemplo, os verbos [comprar] e [vender] seriam vistos como mantendo relações entre pessoas no mundo. Uma palavra como [gato] estaria se referindo a todos os indivíduos da espécie felina. Enquanto um item como [de] só teria existência e sentido em uma relação sentencial. Por essas relações lógicas se configurarem de uma maneira tão ampla, é tarefa de uma semântica lexical distinguir subconjuntos seletivos para cada uma dessas categorias. Isso é possível através da realização de uma análise distribucional padrão, delimitada mais à frente neste trabalho, quando faremos uma distribuição de itens maiores e menores que uma palavra ao subdividi-los entre um conjunto fechado e um conjunto aberto de itens da língua.

Antes disso, poderemos propor uma primeira distribuição abstrata dos itens de uma língua da seguinte forma (PALLETIER and SCHUBERT, 1989):

- a) um conjunto concreto de itens: macaco, homem, garoto, gato, cachorro etc., e esse conjunto estaria incluso em algo como animalidade / mamíferos / animado; esses itens demandariam verbos do tipo: latir / comer / miar / falar etc.;
- b) um outro conjunto concreto de itens: pedra, areia, barro, rocha, cascalho etc., e verbos como: latir / comer / miar / falar não seriam possíveis numa sentença com esses itens,



ou seriam distribuídos numa sentença de uma maneira tal diferente dos itens concretos em (a).

A semântica dos papéis temáticos já delimitou exhaustivamente a relação entre argumentos e verbos, e esse não é nosso propósito aqui, dissemos isso antes. A separação, porém, entre verbos que demandam qualquer tipo de item à sua esquerda como argumento externo como na sentença abaixo:

(42) \_\_\_\_\_ *permanece alegre.* // \_\_\_\_\_ *fica leve.* // \_\_\_\_\_ *está aqui.*

e verbos que selecionariam argumentos específicos que possam preencher adequadamente a posição de argumento externo, como na sentença abaixo:

(43) \_\_\_\_\_ *bebe água.* // \_\_\_\_\_ *come peixe.* // \_\_\_\_\_ *fala bem.*

é uma evidência teoricamente forte das relações sintáticas e semânticas entre itens e uma tentativa de distribuí-los segundo categorias para uma semântica lexical na delimitação de unidades como lexemas.

Resumo: neste capítulo, abrimos longos parênteses com o objetivo de estabelecer pressupostos teóricos que vão nos ajudar a focalizar a entrada lexical como um dado da semântica universal das línguas. Retomamos questões e refizemos rapidamente um percurso para estabelecermos bases em direção a um cercamento do conceito de entrada lexical em termos de unidade lexical e lexema. O capítulo seguinte levanta uma questão: como, a partir da seleção de um lexema, podemos, semanticamente, modular os sentidos em termos de unidade lexical?

## 5 SELEÇÃO E MODULAÇÃO DE SENTIDO

FOCO: a modulação de sentidos e a noção de Entrada Lexical

Um dos problemas básicos da semântica lexical é a aparente multiplicidade de usos semânticos de uma entrada lexical unitária sem diferenças gramaticais. Vejamos o que acontece a respeito disso, com as sentenças que seguem:

(43) *O corisco alumiava no alto e na terra era aquele lençol (1) de chuva que não dava permissão de ninguém ver uma rês... (CL, p. 21)*

(44) *Era aquele lençol (2) bordado que foi vendido mais caro, mas forra camas com mais beleza.*

Os itens estão numerados: lençol (1), e lençol (2). Ao agirmos assim, queremos demarcar distâncias semânticas do tipo: (1) contém traços semânticos não contidos em (2) e vice-versa, apesar das similaridades em termos funcionais, formais, morfológicos e sintáticos arroladas em ambas as sentenças.

Notemos, porém, que podemos encontrar as mesmas investidas gramaticais tanto em (1) quanto em (2), as quais vamos explorar um pouco melhor no parágrafo seguinte.

Traços gramaticais, como apontado acima, no entanto, permanecem os mesmos: em termos categoriais, tanto (1) quanto (2) estão no gênero masculino e no singular; em termos sintáticos, desempenham as mesmas funções sintáticas, ou seja, ocupam a mesma posição em ambas as sentenças, podendo constituir um núcleo para o sintagma de que ele faz parte. O item [lençol], nesses contextos, perde traços semânticos, e ganha outros, ou assume algumas partes de traços que existem potencialmente no léxico, e que são comuns a ambas as ocorrências dessa entrada, tanto na sentença em (43) quanto na sentença em (44) .

Em confronto com sequências como essas, trata-se os itens (1) e (2), de:

- a) Um mesmo item lexical? São itens parciais em termos semânticos? Ou são itens diferentes um do outro?
- b) A noção de entrada lexical pode resolver parcialmente o nosso problema? Se pode resolver, trata-se de uma entrada lexical X em (1) e Y em (2), ou são constituintes de outros fenômenos?

Como o que estamos focalizando aqui é a noção de unidade lexical, vamos testar essas unidades (1) e (2), no quadro abaixo, e responder à questão (a). O quadro apresentado por Rocha (1998, p.78) propõe caracterizar uma palavra, vamos adaptá-lo, contudo, à caracterização da unidade lexical:

	<i>Polissemia</i>	<i>Sinonímia</i>	<i>Homofonia</i>	<i>Conversão</i>
<i>Aspectos</i>	<i>Lençol (1)</i> <i>Lençol (2)</i>	<i>01 / 02</i>	<i>01 / 02</i>	<i>01 / 02</i>
<i>Fonético</i>	<i>Sim</i>	<i>Sim</i>	<i>Sim</i>	<i>Não</i>
<i>Semântico</i>	<i>Parcial</i>	<i>Parcial</i>	<i>Sim</i>	<i>Não</i>
<i>Funcional</i>	<i>Sim</i>	<i>Sim</i>	<i>Sim</i>	<i>Sim</i>
	<i>Itens Lexicais Diferentes</i>			

Quadro 1

Pelo quadro de Rocha, os itens (1) e (2), acertadamente, não seriam as mesmas entradas lexicais, dado que não há uma sinonímia integral entre eles. Para resolver essa aparente contradição, e testemunhar que o quadro de Rocha pode ser aplicado parcialmente à definição de entrada, vamos recorrer à noção de variação, conceito similar ao trabalhado na sociolinguística, para admitir uma aparente e/ou potencial *variação semântica* deduzida nestes termos:

<p>Há uma multiplicidade de usos semânticos de uma entrada lexical, e isso parece ser a regra e não a exceção.</p>
--

Quadro 2

Assim como contextos distintos exigem palavras distintas, contextos diversos exigiriam itens lexicais diversas, expressos por lexemas idênticos. E todos os lexemas se comportam desta maneira: cada um deles pode sofrer variação semântica em termos funcionais quando passa a ser unidade lexical. Segundo Cruse (1992, p. 50): “(...) o significado de uma palavra é, em algum sentido, diferente em todo contexto distinto em que ela ocorra”.

Em termos lexicológicos, no entanto, as palavras podem existir em ou fora de algum contexto possível. Não custa reiterar mais uma vez, em termos contextuais/funcionais, trata-se de uma unidade lexical; já em termos não-contextuais/não-funcionais, trata-se de um lexema.

Podemos estabelecer dois tipos distintos de variação em termos de contribuição semântica que uma entrada traz para sentenças diferentes. Duas formas pelas quais o contexto sentencial de uma palavra pode afetar a sua contribuição semântica para a sentença:

- a) um desses tipos de variação envolve seleções contextuais diferentes a partir de um mesmo item; e,

- b) o outro tipo diz respeito à remissão a contextos diferentes de um mesmo item isoladamente.

Vamos partir dos contextos abaixo para melhor compreender os tipos de variação semântica defendidos para (a) e (b), acima. Começemos a tentativa de compreender o tipo explicitado em (a):

(45) *Conhecia a pessoa de passagem, da cerimônia de bom dia, como-vai-come-passou.*

(46) *Conhecia a pessoa dela de passagem, da cerimônia de bom dia, como-vai-come-passou.*

(47) *Conhecia a pessoa dele de passagem, da cerimônia de bom dia, como-vai-come-passou. (CL. p. 22)*

Concentremo-nos no uso do item lexical [*pessoa*]. Do contexto em (45), podem derivar os contextos em (46) e (47). A variação contextual em (46) e (47) é plausível pois o item isolado [*pessoa*] pode possuir leitura semântica e pragmaticamente viáveis para esses contextos. Ele comporta, apesar de ser uma palavra de gênero único, a possibilidade de referência a homem ou a mulher, e, portanto, tem dupla remissão semântica.

Mas, tanto (45), (46) e (47) podem funcionar como comunicação satisfatória, dado que se confia no falante, em sua capacidade de cooperar semanticamente, já que ele procura cumprir o contrato semântico, e quer ignorar o sexo do referente ou pode cobri-lo mentalmente sem prejuízos em termos semânticos.

Mas não se pode ignorar que nas três ocorrências esse item *é ele* e *não é ele*, e que os falantes atribuem a esse item um sentido geral que cobre as “particularidades” ou possibilidades específicas em (46) e (47).

Vamos partir agora para o tipo explicitado em (b):

(48) *Ele finalmente encontrou o banco.*

O item [*banco*], que aparece na sentença em (48), também pode sofrer interpretação variada, mas atribui uma ambiguidade a essa sentença de que [*pessoa*] em (45) não é capaz. É uma sutileza semântica que pode ter repercussões estilísticas e surpreendentemente sígnicas para ser ignoradas por uma descrição que se pretenda exaustiva em termos semânticos.

Uma entrada isolada [*banco*], em (48), pode remeter os falantes para contextos variados. Já uma palavra isolada [*pessoa*], em (45), pode *ou não* remeter os falantes para contextos variados – não se exige necessariamente que essa remissão seja feita, pois o contexto linguisticamente montado numa sentença por si só já é capaz de estabelecer as bases semânticas para cumprir as “cláusulas” do “acordo semântico”.

Devemos aqui fazer uma distinção entre traços semânticos que se quer *gerais* para os itens, é o caso de [*pessoa*], em (45): o traço é geral pois esse item pode ser usado tanto para incluir pessoas do sexo feminino quanto pessoas do sexo masculino. Já traços semânticos que geram *ambiguidade*, e presentes em um mesmo item, jamais poderiam incluir ao mesmo tempo mais de uma interpretação: os falantes devem escolher uma interpretação para o contexto viável em que operam, por exemplo:

(49) *Ele finalmente encontrou o banco e se sentou.*

(50) *Ele finalmente encontrou o banco e pagou a conta.*

Enquanto os contextos (49) e (50) são excludentes entre si, os contextos (47) e (48) são complementares, podendo ser delimitados por um contexto mais amplo, como em (9), abaixo:

(51) *Conhecia a pessoa dele e/ou a pessoa dela de passagem, da cerimônia de bom dia, como-vai-come-passou.*

Se fizessem parte de um dicionário mental em termos semânticos, seus lexemas seriam distribuídos em entradas, no mínimo muito diferentes entre si: [*pessoa*] inclui tanto seres humanos do sexo masculino quanto feminino, no entanto, realidades contextuais similares (seres humanos); [*banco*] refere-se a realidades contextuais diversas: [*banco, o móvel*] e [*banco, a instituição financeira*].

Agora, vamos examinar com mais detalhes diferentes maneiras pelas quais cada contexto exerce influência restrita em relação aos sentidos associados a entradas que ocorrem dentro desses contextos.

Há duas maneiras fundamentais pelas quais a contribuição semântica efetiva de um item lexical pode variar sob influência de diferentes contextos, e cada contexto enfatiza traços semânticos e obscurece ou suprime outros. Vamos, ainda, utilizar os pressupostos teóricos de Cruse (1992, p. 52) para problematizar as relações semânticas de CL: Ao “efeito de um contexto incluindo uma unidade lexical, chamaremos de *modulação*. Modulação é a variação dentro de um significado, que é amplamente contínua e fluida por natureza”.

A segunda maneira de variação semântica diz respeito à ativação de variados contextos, que remete a variados sentidos associados a uma unidade lexical ambígua, como foi o caso problematizado em (45), (46) e (47). A que Cruse (1992) se refere como seleção contextual de sentidos.

Esses dois tipos geralmente agem de forma contínua na ativação de sentido imaginada pelos falantes. O que quer dizer que a ativação de sentido é algo sujeito à modulação por contextos manipulados pela seleção. Vamos focalizar um pouco mais a modulação. Há duas formas centrais que estabelecem o *status* dos traços semânticos em termos de sua modulação, em relação à aquisição e à manipulação desses traços. Acompanhemos a descrição no quadro abaixo:

MODULAÇÃO DE TRAÇOS SEMÂNTICOS	
AQUISIÇÃO DE TRAÇOS SEMÂNTICOS	MANIPULAÇÃO DE TRAÇOS SEMÂNTICOS
PROGRESSIVOS	POR
E	ALÇAMENTO
REGRESSIVOS	E
	ANTECIPAÇÃO

Quadro 3

Essa divisão, adaptada da proposta de Cruse (*op.cit.*), restringe ou amplia, de certa maneira, o escopo semântico de um item lexical em função numa sentença. Vamos considerar o conjunto de sentenças abaixo e entender como as noções de traços progressivos e regressivos são delineadas pela língua:

(52) *“Inventei compromisso de mulata teúda e manteúda de propósito, quando nenhuma gente sabe que nenhuma cara bonita prende em cativoiro homem como o neto do velho Simeão.”* (CL, p. 109)

(53) *Inventei compromisso de mulata teúda e manteúda de propósito, quando nenhuma gente masculina sabe que nenhuma cara bonita prende em cativoiro homem como o neto do velho Simeão.*

O foco aqui é o item lexical [*gente*] que aparece em (52), e em (53) modificado pela palavra restritiva [*masculina*]. Esperam-se traços gerais em (52), traços gerais aqui incluindo ambos os sexos. Mas em (53), embora não tenha sido modificado o seu gênero gramatical, há uma restrição considerável de escopo desse item.

Esse conjunto de sentenças é suficiente para explicitar a modulação, progressiva e regressivamente imaginada pelos falantes ao retirar de seu dicionário mental um lexema e

transformá-lo em unidade lexical. Em outras palavras, há uma restrição em termos de um contrato semântico<sup>16</sup> que os itens assumem em função numa sentença.

É talvez óbvio, mas inegável, e imprescindível, que em função numa sentença os itens perdem traços, ou ganham outros, apesar de manter um núcleo estável variando de maneira suficiente para fins semânticos. Os itens estão entrelaçados de tal modo a outros que eles chegam a perder uma imaginada identidade de sentido unitária ou fixa.

Um dos aspectos relevantes da modulação de sentido é a **promoção** de traços semânticos – um traço como geral em (52). Ou a **regressão** desses traços, a restrição de sentido a que um lexema como [*gente*] é submetida em (53).

Vamos pontuar a essas alturas de nossa argumentação que não estamos fixando exaustivamente os traços associados a cada item lexical, nem tampouco fazendo um mapeamento de todos os traços consideráveis para um item em particular, isso seria uma tarefa hercúlea e/ou quixotesca pois traços elencados seriam contestados ou muitos deles poderiam faltar em nossa descrição. O nosso intuito é estabelecer padrões gerais de funcionamento e consideração desses traços para o propósito que aqui nos interessa.

Um outro efeito da modulação de sentido de uma unidade lexical envolve a relação entre **alçamento** e **antecipação** de traços semânticos. Diferentes tipos de traços podem ser aferidos dessa forma. E isso é algo tão natural em termos de construção e/ou contrato semântico entre os usuários de uma língua que muitas vezes não nos damos conta das sutilezas suficientes e sofisticadas da construção de sentido. Vamos observar o comportamento sígnico dos itens que ocorrem nas sentenças abaixo:

(54) “**Contei nos dedos – segunda, terça, quarta, quinta, sexta, sábado. Calhava em data de sete...**” (CL, p. 51)

(55) **Contei nos dedos – os dias da semana. Calhava em data de sete.**

Na sentença em (54), há um alçamento bem claro das partes componentes dos dias da semana. Já na (55), podemos identificar uma antecipação, a ser preenchida pelas inferências dos usuários da língua: algo em torno do esquema abaixo:

(56) (...) **os dias da semana** ( \_\_\_\_\_).

---

<sup>16</sup> Como antecipado em nota anterior, há um contrato semântico entre os usuários de uma língua quanto ao menos eles reconheçam um traço em comum que faz parte de um item lexical; ou quando eles reconheçam contextos possíveis para os quais ao menos um traço de um item possa se encaixar nesse contexto.

Em que o espaço ( \_\_\_\_\_ ) pode ser preenchido pelos usuários da língua. É ponderável, contudo, que tanto a sentença (54) quanto a (55) cabem em contextos discursivos idênticos. Mas a que mais se revela é a (54), dada a focalização proposta pelo desdobramento detalhado de um conjunto – os dias da semana.

As sentenças em (54) e (55) fazem referência a uma mesma realidade, não existem dias da semana que não sejam aqueles demonstrados em (54); assim como em (55), os dias da semana não poderiam ser diferentes dos elencados em (54).

Em relação aos sentidos modulados contextualmente, podemos ainda afirmar que uma entrada lexical pode ser diferentemente apreendida em cada contexto distinto em que aparece. Por exemplo, em:

(57) *Contei nos dedos os dias da semana – segunda, terça, quarta, quinta, sexta, sábado. Calhava em data de sete...*

O item lexical complexo [*dias da semana*] pode sofrer modulações (que aparecem em (54), (55) e (57)). Em cada uma dessas modulações, contudo, ele pode ganhar e perder traços; pode ter esses traços suprimidos ou pode encaminhar antecipação ou alçamento de traços semânticos.

Estamos argumentado aqui os efeitos de um contexto sobre o sentido de um item lexical em particular. Podemos dizer que ele tem seus traços semânticos modificados ao sabor do jogo semântico entre os usuários de uma língua como o português.

Mas o contexto também pode atuar de maneira tão selvagem sobre um item lexical em particular que, aquilo que outrora tem uma estabilidade semântica precária pode ser ambíguo ou radicalmente modificado em relação ao contexto de que venha a fazer parte. As sentenças retomadas podem nos ajudar a visualizar a atuação de um contexto de radicalização selvagem de sentido:

(48) - *Ele finalmente encontrou o banco.*

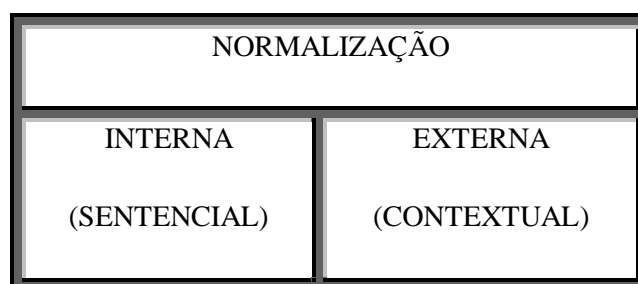
(17) - *Aquilo é tábua de passar roupa. Moça para o primo tem que ter coxão fornido, capaz de agüentar os repuxos. (CL, p. 17)*

Em questão, estamos nos referindo aos itens [*banco*] e [*tábua de passar roupa*]. Esses itens podem assumir significados tão distintos e inesperados, a partir da manipulação contextual, que é tentador assimilá-los como opacos semanticamente. Ou então poderíamos pensar de uma outra maneira: há, por um lado, uma seleção semântica, na qual o falante elegeria o sentido que convém ao item; ou uma seleção contextual, na qual o falante selecionaria o contexto adequado onde caberia o sentido construído em negociação semântica num jogo de cena comunicativo.



Em outros termos, um falante assumiria o que um produtor de uma sentença quer dizer e, então, escolheria (semântica zero) um contexto para que a sentença seja interpretada: há um lugar de onde se fala e um lugar para o qual um item lexical pode ser adequado para um contexto.

Poderíamos, para continuar o nosso raciocínio. Vamos supor dois tipos de normalização de uma sentença para fins de comunicação: uma normalização interna (sentencial) e uma normalização externa (contextual), e os falantes de uma língua estão tão conscientes sobre isso que logo pressupõem relações possíveis para satisfazer os processos semânticos de construção de sentido.



Quadro 4

Como conseguimos nos comunicar com itens ambíguos, como os que aparecem em sentenças acima? Por que a ambiguidade é uma estratégia de sentido por vezes condenada pela gramática tradicional, mas é também um recurso de expressão tão rico que jamais pode ser abandonado em descrições linguísticas sérias? Essas questões podem ser cercadas por assertivas que deduzimos nos seguintes termos:

- a seleção adequada de itens pode auxiliar os falantes na efetiva comunicação por línguas naturais; como também,
- a seleção contextual pode desfazer a ambiguidade.

Isso resulta no seguinte:

- todo item de uma língua pode assumir relações ambíguas, desfeitas por contextos adequados; e,
- há contextos que podem ambigüizar itens ou sentenças.

Pensando assim, podemos assumir que a ambigüidade, condenada por professores de gramáticas tradicionais, são recursos ricos e expressivos de uma língua, que possuem poder estilístico efetivo. E toda sentença ambígüa assume uma interpretação relevante de primeiro plano capaz de desfazer a ambigüidade. Retomemos a sentença em (48):

(48) *Ele finalmente encontrou o banco.*

Não é preciso ir muito longe em termos argumentativos para desfazer essa ambiguidade, ressaltando o poder que os contextos podem assumir para isso. Reservaremos um capítulo em que investiremos com mais acuidade nas relações entre entradas lexicais e ambiguidade. No próximo item, vamos problematizar como os falantes conseguem entradas lexicais.

## 5.1 CONSEGUINDO ENTRADAS LEXICAIS

A uma sequência fonológica se acopla um jogo cênico de sentido para que possa servir de base comunicativa e discursiva num mundo linguístico e extra-linguístico.

Vamos começar entendendo como crianças possuem um domínio excepcional de sentido de palavras que parte de uma sequência de itens, digamos, não-semântica, para algo a que vamos chamar de sequência de conteúdo útil para a comunicação. Para tanto, é necessário a estas alturas algumas considerações a respeito da aquisição de primeira língua pelas crianças. Tarefa para os próximos parágrafos.

Vamos considerar quais princípios podemos propor para uma teoria de aquisição da língua materna pelas crianças, e relacionar isso com os propósitos sobre a aquisição da entrada lexical. E vamos ter como evidência que a imitação – no sentido de que as crianças, para aprenderem a falar, imitam a fala de outras crianças ou a dos adultos que a cercam – está fora das estratégias de aquisição.

Uma teoria de aquisição da linguagem diz respeito à questão de como as crianças adquirem gramáticas de suas línguas nativas.

As crianças geralmente produzem suas primeiras entradas reconhecíveis (como “mamã” ou papá”) por volta dos primeiros doze meses de idade. Nos seis meses seguintes em diante, há uma pequena aparência evidente de desenvolvimento gramatical em sua produção de fala, embora a produtividade do vocabulário tipicamente cresça em torno de cinco palavras por mês até que alcance em torno de trinta palavras por volta dos dezoito meses. (RADFORD, 2004)

De um estágio inicial de uma entrada única até a completa aquisição da gramática de suas línguas, as primeiras tentativas de frases das crianças compreendem entradas soltas, faladas isoladamente. Por exemplo, uma criança pode falar “maçã” quando precisa dizer “quero uma maçã”, ou diz “cima” quando quer subir em cima de uma cadeira.

Durante o estágio de entradas soltas, é difícil encontrar evidências claras de aquisição da gramática – já que a gramática é uma complexidade que envolve o domínio de regras de encadeamento de palavras, noções sofisticadas de significados, uso de tempos verbais mais

abstratos, concordâncias do tipo “o pires”, “os pires” etc –, obviamente, isso não faz parte desse estágio inicial ainda, já que o nível de produtividade da linguagem infantil não consegue combinar uma entrada com uma outra para formar frases de três ou mais entradas como: “mamãe, eu quero uma maçã”.

Por volta dos dezoito meses (embora com considerável variação de criança para criança), descobrimos os primeiros sinais visíveis da aquisição da gramática: as crianças começam a fazer uso produtivo das flexões (por exemplo, usando o plural de uma entrada como “mesa” (mesas), ou flexões verbais como “eu como”, “você come”, e, de forma similar, começam a produzir frases elementares contendo mais de duas entradas, como as que pontuamos acima). Deste ponto em diante, há uma rápida expansão em seu desenvolvimento gramatical até que, por volta dos trinta meses, elas tenham adquirido tipicamente mais flexões e mais construções gramaticais centrais usadas no português, e até que estejam aptas a produzir frases como aquelas que os adultos produzem, tais como: “Onde mamãe foi”, “O que papai está fazendo?”, “Vamos comprar doces, papai?” etc (embora erros sintáticos e morfológicos ocasionais persistam até a idade de quatro ou cinco anos ou mais – tais como: “eu trazi o vaso, vovó”).

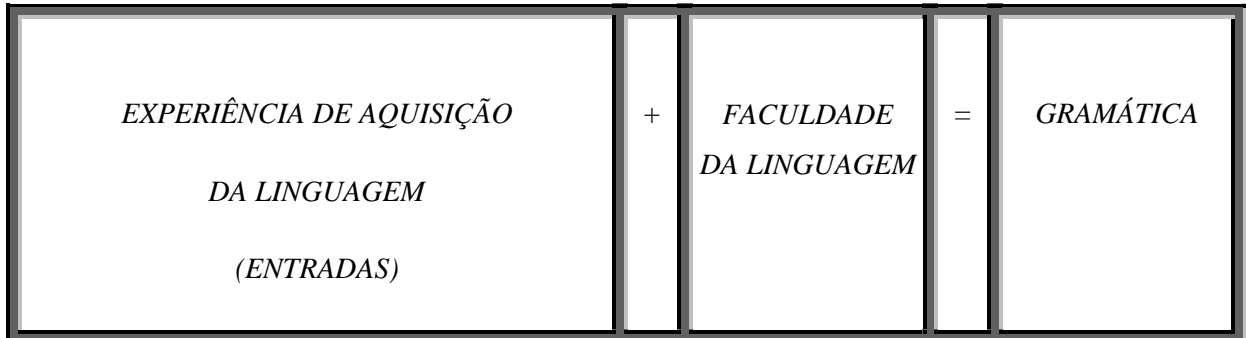
Dessa forma, o fenômeno central que qualquer teoria de aquisição da linguagem deve buscar é este: como é que, depois de um arrastado longo período de muitos meses, não há sinais óbvios de desenvolvimento gramatical e por volta dos dezoito meses há um súbito jorro de frases contendo mais de duas entradas disparadas que começa a emergir, e um desenvolvimento fenomenal da gramática então acontece ao longo dos próximos doze meses?

A *uniformidade* e (uma vez que o jorro dispara) a *rapidez* no padrão de desenvolvimento linguístico das crianças são fatos centrais que uma teoria de aquisição da linguagem deve buscar explicar. Mas como?

Chomsky (1972) sustenta que a explicação mais plausível para a uniformidade e a rapidez de aquisição da língua materna é propor que o curso da aquisição é determinado por uma faculdade de linguagem de que os seres humanos são dotados biologicamente. Por faculdade da linguagem entende-se um programa de aquisição – que ele toma de empréstimo como metáfora de um programa de computador – dentro do cérebro, que proporciona às crianças um algoritmo transmitido geneticamente, isto é, uma configuração de procedimentos, para que elas desenvolvam uma gramática, com base em suas experiências de linguagem (ou

seja, tocadas pelos dados – as entradas que elas ouvem no meio-ambiente social onde estão imersas – do mundo).

A forma com que Chomsky visualiza o processo de aquisição pode ser representada neste quadro abaixo:



Quadro 5

As crianças, no processo de aquisição, observarão as pessoas ao redor delas usando a língua, e o conjunto de expressões da língua que uma criança ouve (assim como os contextos em que a língua é usada) no percurso de aquisição constitui sua experiência linguística.

A experiência – os dados do mundo real, ou seja, as entradas que ela ouve dos adultos e de outras crianças – que uma criança nos estágios iniciais de aquisição da língua materna cultiva serve como *input* (entrada) para a faculdade de linguagem (espécie de órgão contido no cérebro dos humanos que desencadeia a aquisição da língua materna), que proporciona à criança procedimentos de análise subconsciente dos dados, e elabora a gramática da língua sendo adquirida.

Dessa forma, o *input* para a faculdade de linguagem é a experiência com os dados, e o *output* (saída) da faculdade de linguagem é a gramática de uma língua em processo de aquisição.

A hipótese de que o curso de aquisição da linguagem é determinado por uma faculdade de linguagem inata é conhecida na literatura como hipótese inatista. Em contraposição a hipóteses que pressupõem ser a língua adquirida por imitação, no contato com a cultura, propostas por teorias do comportamento (como o behaviorismo).

Chomsky avalia que a habilidade para falar e adquirir uma língua é exclusiva da espécie humana, e que as línguas incorporam princípios que são exclusivos dos humanos, refletindo a natureza da mente humana.

Traduzo alguns excertos do livro “Language and Mind” (1972) para vermos como ele se pronuncia a respeito da hipótese inatista de aquisição:

Todas as evidências que temos parecem sustentar o ponto de vista de que a habilidade de aquisição e uso de uma língua é uma capacidade específica da espécie humana, que há princípios profundos e restritivos que determinam a natureza da

linguagem humana enraizados no caráter específico da mente humana. (CHOMSKY, 1972 p.102)<sup>17</sup>

Além disso, ele nota, a aquisição da língua materna é uma habilidade que todos os seres humanos possuem, independente inteiramente de sua inteligência geral – nesses termos, ele questiona o envolvimento da inteligência no processo de aquisição, como propõe teorias construtivistas de base piagetiana.

Ele defende que a inteligência está em segundo plano no processo de aquisição nestes termos:

Mesmo com níveis baixos de inteligência, ou níveis leves de patologias (uma criança adquire uma língua), encontramos (nas crianças) um comando de linguagem que é totalmente irrealizável em macacos, em outros aspectos, (o comando) supera determinados níveis de patologias na aquisição como também nas atividades para resolver problemas utilizando uma língua, assim como outros comportamentos adaptativos. (CHOMSKY, 1972 p.10)<sup>18</sup>

Além disso, a aparente uniformidade nos tipos de gramáticas desenvolvidas por diferentes falantes da mesma língua sugere que as crianças têm uma orientação genética na tarefa de construir uma gramática de sua língua nativa.

Quando pressupomos que a uniformidade é aparente para os mesmos falantes da língua portuguesa, estamos defendendo aqui que circula entre os brasileiros nativos da língua portuguesa diferentes tipos de falar o português, refletindo diferentes manifestações da gramática do português. Vide, o “jeito” baiano de falar, em contraposição ao “jeito” mineiro, ao “jeito” de comunidades isoladas etc. Em relação a isso, Chomsky assim se expressa:

Sabemos que as gramáticas são fatos em construção e que variam ligeiramente entre falantes de uma mesma língua, apesar de haver variações mais amplas não somente em termos de inteligência como também em termos das condições sob as quais uma língua é adquirida. (CHOMSKY, 1972, p.79)<sup>19</sup>

Ademais, a rapidez da aquisição (uma vez que o jorro das frases tem início) também evidencia uma orientação genética na construção da gramática. Mais uma vez, ele com a palavra: “(...) caso contrário é impossível explicar como as crianças começam a construir sua

---

<sup>17</sup> Whatever evidence we do have seems to me to support the view that the ability to acquire and use language is a species-specific human capacity, that there are very deep and restrictive principles that determine the nature of human language and are rooted in the specific character of the human mind.

<sup>18</sup> Even at low levels of intelligence, at pathological levels, we find a command of language that is totally unattainable by an ape that may, in other respects, surpass a human imbecile in problem-solving activity and other adaptive behaviour.

<sup>19</sup> We know that the grammars that are in fact constructed vary only slightly among speakers of the same language, despite wide variations not only in intelligence but also in the conditions under which language is acquired.

gramática num curto espaço de tempo com base em sua experiência tão limitada de acesso aos dados”. (CHOMSKY, 1972, p. 113)<sup>20</sup>.

Uma criança conhece tão profundamente a sua língua em um período tão curto de tempo, e isso desbanca as propostas de aprendizado baseadas na imitação e no comportamento, como preconizam teorias não inatistas. Se os pressupostos das teorias não inatistas fossem levados em consideração, um ser humano começaria a falar quando chegasse à idade de 80 anos! ,dadas as sofisticacões na construção de frases de que uma criança já é capaz de produzir. Por imitação, portanto, não será possível saber tanto da língua em tão pouco tempo.

O que faz com que a uniformidade e a rapidez do processo de aquisição sejam dignos de nota é o fato de que a experiência linguística é geralmente muito imperfeita, já que se baseia na fala dos adultos, que pode ser uma transposição pobre de sua competência como falantes.

Por exemplo, os adultos dizem frases incompletas, constroem frases sem sujeito (uma mãe pode estar conversando algo com uma amiga, e diz “chegou ontem”, e a criança em processo de aquisição só ouve isso, por exemplo). Outra coisa, como um bebê de poucos meses vai saber a diferença entre o latido de um cão e uma palavra? Nesses termos, a imitação está totalmente fora do processo.

Se muitas entradas (*input*) de fala que as crianças recebem são imperfeitas (ou seja, como vimos, elas ouvem muitos erros, falas “quebradas”, frases pela metade, sons inaudíveis, murmúrios, e sons muito diferentes uns dos outros), como então elas conseguem utilizar essa experiência deficiente para desenvolver sua gramática que especifica como formar sentenças no processo de aquisição da língua materna? Traduzo excertos retirados da obra de Chomsky, dessa vez de “Interview with S. Hamshire”, in *The Listener* (1968), em que ele delinea ideias em defesa bem sistematizada da hipótese inatista (suposição de que a linguagem é herdada geneticamente). Para responder a essa questão, ele investe na seguinte analogia:

Descartes questiona: como conseguimos ver um tipo de figura irregular desenhada na nossa frente e sabemos se tratar de um triângulo? Ele observa, muito corretamente, que há uma disparidade entre os dados apresentados a nós e a percepção que deles conseguimos obter. Ele argumenta – e avalio, com muita plausibilidade – que vemos a figura como um triângulo porque há algo da natureza de nossas mentes que faz com que a imagem de um triângulo seja facilmente construível por nossas mentes. (CHOMSKY, 1968, p.687)<sup>21</sup>

<sup>20</sup> Otherwise is impossible to explain how children come to construct grammars... under the given conditions of time and access to data.

<sup>21</sup> Descartes asks: how is it when we see a sort of irregular figure drawn in front of us we see it as a triangle? He observes, quite correctly, that there's a disparity between the data presented to us and the percept that we

A implicação óbvia é que, como há muitas maneiras com que somos geneticamente predispostos a analisar formas (entretanto, irregulares) como tendo propriedades geométricas, assim também somos geneticamente predispostos a analisar sentenças (entretanto, não gramaticais) como tendo propriedades gramaticais específicas.

Um outro argumento que Chomsky usa para sustentar a hipótese inatista se relaciona ao fato de que a aquisição da linguagem é inteiramente subconsciente, e uma atividade involuntária – ou seja, a criança não tem escolha: vai falar uma língua sem que tenha opção de não querer aprendê-la – no sentido de que não tem consciência se quer adquirir ou não a língua nativa – embora, mais tarde, podemos escolher se queremos ou não montar bicicleta ou dirigir um carro.

A aquisição de uma língua é também uma atividade em nada monitorada – no sentido de que os pais não ensinam uma criança a falar: ela aprende por si mesma. Chomsky, nesses termos, se pronuncia: “As crianças adquirem... a língua materna com invejável sucesso, apesar de nenhum cuidado especial dos pais para que ela aprenda, e nenhuma atenção especial seja necessária para seu progresso”<sup>22</sup>. (CHOMSKY, 1965, p. 2001). Isso implica que não se aprende a ter uma língua nativa, nem tampouco uma criança aprende a ter braços e pernas – ela os tem, simplesmente. A habilidade de possuir uma língua nativa é parte de nosso legado genético – assim como a habilidade de aprendermos a andar.

Os estudos de aquisição da linguagem se valem de suportes empíricos da hipótese inatista. Há um *período crítico* para a aquisição da sintaxe, no sentido de que as crianças que aprendem uma outra língua além da língua materna antes da puberdade, geralmente alcançam um grau notável de domínio dessa segunda língua, enquanto que aquelas que aprendem uma segunda língua depois dos nove ou dez anos de idade raramente dominam-na com a mesma competência sintática de um falante nativo.

O que sugere que a aquisição da sintaxe é determinada por um programa de aquisição da língua inato que é, com efeito, totalmente preenchido no início da puberdade, ou seja, a partir da puberdade, os espaços sintáticos vazios na mente já foram preenchidos pela língua materna, e não há mais lugares a ser preenchidos por uma outra língua. Nesses termos, como

---

construct. And he argues, I think quite plausibly, that we see the figure as a triangle because there's something about the nature of our minds which makes the image of a triangle easily constructible by the mind.

<sup>22</sup> Children acquire... languages quite successfully even though no special care is taken to teach them and no special attention is given to their progress.

bem observou Pinker (1994), a sintaxe pode ser vista também como um órgão, ou um módulo da mente humana herdado da evolução.

Temos observado que a hipótese inatista fixa a faculdade da linguagem como um dispositivo capaz de acionar a aquisição da língua pelas crianças. Cabe, portanto, nos questionarmos em relação a sua natureza e o que pode ser incorporado pela faculdade da linguagem.

A noção de faculdade da linguagem não é recente nos estudos linguísticos, e nem tampouco foi proposta por Chomsky. Em “Ensaio sobre a origem das línguas”, Jean-Jacques Rousseau (1982 [1750]) concebe que reside no cérebro humano algo como se fosse um órgão responsável pela linguagem, que levaria as palavras e frases para a boca, para serem pronunciadas pelas pessoas.

Juntando essa intuição com uma outra, do cientista alemão Humboldt (1994 [1817], p.37), reconhecido como o primeiro linguista europeu, a de que “a mesma gramática está na cabeça de todas as pessoas”, podemos propor que a natureza da faculdade da linguagem é sua capacidade de ser universal, e ela incorpora a possibilidade de haver um momento em que, para todos os bebês, antes de que o processo de aquisição seja disparado, eles possam aprender quaisquer línguas.

Sendo assim, um bebê que nasceu no Japão, vindo para o Brasil ainda nos primeiros meses, logicamente irá aprender o português, e vice-versa. A hipótese inatista de aquisição não encontra rivais na literatura acadêmica que lhe apresente pressupostos tão bem fundamentados quanto os que ela defende. Uma criança aprenderia a falar sem que houvesse contato com outros falantes, ou seja, com a cultura? Obviamente que não. Mas assim que o contato se efetua, uma criança não precisa de mais nada para falar a língua que fala, e começar a conseguir as entradas que a acompanharão por muito tempo em sua vida adulta.

Primeiramente, as crianças aprendem um inventário de entradas com função e conteúdo. Cada criança descobre e cria o seu dicionário – seu conjunto de lexemas. Cada uma delas, de um modo um pouco diferente uma das outras, vai pouco a pouco anexando sentidos a uma sequência aparentemente nua semanticamente. Todas elas, contudo, possuem uma tendência natural que funciona na anexação de conteúdos a uma sequência inicialmente desprovida de base semântica.

Essa tendência, à primeira vista, é influenciada pelos adultos e pelas crianças que cercam a criança em processo de construção de seu dicionário mental. Essa proeza é possível graças a uma capacidade inerente à infância de perceber os sentidos já realizados, consumados social e historicamente. Mas, atenção, os sentidos não são simplesmente



captados entrada a entrada: os sentidos são gerativos, ou seja, as crianças não adquirem o sentido de itens isoladamente, um a um, item a item, mas a partir de um *parse*, ou seja, de uma análise mental que funciona em cadeia e em relações mentais herdadas e inatas.

Os sentidos captados e distribuídos pelas crianças para a comunicação, diferente do que pensa Heidi (2005), não dependerão inteiramente das crianças e dos adultos que a cercam, ou do mundo linguístico e extra-linguístico no qual ela está inserida, mas da capacidade de composição mental das crianças.

É claro que a capacidade para a construção de um dicionário é histórica e social, advém da comunidade de fala na qual a criança está inserida. Mas perguntas simples podem fundamentar a noção mentalista que defendemos aqui: por que crianças imersas numa mesma comunidade, ou irmãos vivendo e compartilhando o dicionário de uma mesma família, não internalizam a mesma lista de itens lexicais? Por que a lista das gerações anteriores não permanece a mesma para as gerações futuras? Haveria uma proto-lista que sobreviveria ao tempo e permaneceria a mesma para a aquisição?

Bom, para a terceira questão, é claro que não. A lista-inventário de uma língua muda de ano a ano, de décadas a décadas, de séculos a séculos. É por isso que temos de tempos em tempos um dicionário novo sendo lançado, ou temos dicionários especializados. A cada momento histórico uma variedade de novas entradas é inventariada. É por isso que uma região se distingue linguisticamente de outra, ou que há conservadores e vanguardas, pessoas cuja lista consta de mais itens que outra, dicionários maiores e menores.

No processo de aquisição do léxico pelas crianças, dão-se, ressaltadas proporções ideológicas, temáticas, regionais, sociais e comunicacionais, diferenças de montagem de listas mentais de itens.

Quem decide, assim, que uma lista se torne defasada ou mais ampla? Que uma palavra se torne mais popular que outra, que um item entre e saia de moda? Certamente não são linguistas ou filólogos – dicionários e filólogos apenas registram o que aconteceu em termos lexicais, e nunca o que vai acontecer. Ou seja, eles não têm o poder de criar ou impor itens novos: pusilânimes em termos lexicais, dicionários e filólogos, apenas observam a vida lexical acontecer, e apenas, arqueologicamente, escavam o fosso lexical para montar as suas listas.

Raramente alguém consegue deliberadamente cunhar um novo item como uma nova entrada lexical. É um equívoco anti-científico linguisticamente, como fez uma grande revista

de circulação nacional quando da morte do colunista Ibrahim Sued, atribuir a ele a criação do item “pantera”, para designar uma moça sexy. O item, com total probabilidade, já circulava em certa comunidade, e foi usado por Ibrahim em sua famosa coluna: em termos de léxico dado, uma comunidade conta mais do que alguém isoladamente. O que podemos cientificamente deduzir, no entanto, é que novas palavras são criadas de itens pré-existentes, num jogo constante de eterna reciclagem.

Um grande problema enfrentado pelos linguistas na discussão de como um novo item aparece para uma língua é decidir o que conta de fato quando genuinamente surge uma nova palavra, já que uma palavra surge de uma outra, já que o processo de reciclagem é constante e base de formação.

Nessas alturas, podemos introduzir a noção de formação retroativa. Para deduzir que temos, diante de nossos ouvidos, sentidos e olhos, uma nova palavra, a morfologia sozinha não basta: devemos nos cercar da noção de léxico gerativo, da semântica, de relações extra-linguísticas, e da sintaxe. Isso porque o sentido de uma palavra em particular pode mudar através dos tempos, mas a sua forma permanecer a mesma. Ou o contrário, a forma pode mudar para adquirir novos sentidos, ou sentidos antigos.

De outra forma, a conexão entre o sentido original e o sentido moderno de um mesmo item de mesma forma mudou significativamente a tal ponto que os dicionários registram esse novo item como palavras distintas.

As línguas adquirem novas entradas constantemente, mas é muito difícil dizer quando essas novas entradas apareceram. A gradação semântica de um item como terror, por exemplo, é tão surpreendente, que mal podemos dizer se se trata de um item novo a cada contexto, ou de uma entrada com traços semânticos sutis em variação, ou se cada entrada de verbete de dicionário pressupõe uma nova criação. Senão, vejamos:

58) Essa garota é um terror.

59) Márcio assistiu a um filme de terror.

60) O terror abateu a cidade.

No contexto (58) a garota jamais poderia assumir a qualidade dos contextos (59) e (60), assim como os contextos (59) e (60) não se recobrem totalmente para a entrada terror.

Nesses termos, levantamos duas questões:

a) quando a língua adquire uma nova entrada ou um novo lexema?, e

b) ou é uma mesma entrada com sentido mudado?

Vamos assumir, então, com Heidi (2005), que uma nova dada sequência sonora, com um novo sentido, que pode ser variação sutil de um sentido já posto, é uma nova entrada na

língua, nos termos em que uma mudança semântica é radical o suficiente na medida em que uma sequência sonora que nos é familiar assume o *status* de uma nova palavra na língua. Vamos continuar observando algumas coisas no português que os falantes fazem para criar novas entradas em termos de formação retroativa.

## 5.2 FORMAÇÕES RETROATIVAS

Entradas particulares – conexões som-sentido particulares – são recriadas de uma outra maneira na mente de cada criança que os assimila. Esse processo de aquisição, por ser incrivelmente preciso, é possível por uma atividade de incessante deslizamento semântico – entendendo como deslizamento semântico a atividade eficiente que preserva o sentido suficientemente para a comunidade na qual a criança está inserida.

Uma criança adquire uma entrada que nunca ouviu antes das pessoas que a cercam, dadas circunstâncias favoráveis, pois tem a mente estimulada, como pontuamos acima por um *input* de formações retroativas, ou seja, por ter sido tocada por formações anteriores.

Uma maneira comum de novas entradas entrarem para a lista das crianças, mesmo a entrada sendo totalmente nova para elas, é feito via processos que criaram entradas: re-análise e acoplagem semântica de itens. Se a re-análise de um item como [invencion-], ou de um item como [-eiro], e a acoplagem desses itens como na sequência [invencioneiro] é feita suficiente e eficientemente para funcionar numa cena comunicativa, e que seja possível recriar essa sequência na mente de um ouvinte em termos de conteúdo, então temos uma nova entrada, portanto, um novo lexema.

Nesses termos, itens que semanticamente funcionam isoladamente (palavras ou não) na mente de um adulto, podem vir a se tornar uma nova palavra para a comunidade de fala por compartilhar a mente tanto da criança em processo de aquisição quanto do adulto em processo de interação com a criança.

De forma similar, palavras formadas de múltiplas partes, como é o caso da acoplagem de itens como [caneta] e [chicote] num item único [caneta-chicote] funcionam como uma entrada única, apesar de o item ser, em termos de composição, um item complexo. Nesses termos, entradas que nunca existiram antes podem ganhar vida e assumir existência real.

Segundo Heidi (2005), re-análise e acoplagem de itens dão-se de duas formas: formações retroativas (*back formations*) e etimologias populares (*folk etymologies*). Os itens

ao longo deste texto, oriundos do romance o CL, e de outros, são produzidos, em termos mentalistas a partir do critério “back formations”, e em termos descritivos, a partir do critério “folk etymologies”.

Segundo Heidi, o segundo caso ocorre na medida em que alguém dá o seu melhor para que um item estranho ou não familiar ganhe força semântica e possa ser entendido de uma forma precisa o mais próximo possível de entradas semanticamente já presentes numa comunidade de fala – já que toda nova formação precisa estar próxima semanticamente de uma formação já existente para que o seu conteúdo seja registrado na mente de uma pessoa que nunca tenha ouvido essa nova formação anteriormente. Sabe-se o que significa [invençioneiro], mesmo nunca tendo ouvido esse item antes, por a mente ser capaz de acoplagens e re-análises semânticas.

Em termos de etimologias populares, por exemplo, o falante que diz [*corgo*] por ter se “enganado” ao não ouvir “direito” *córrego*, faz um deslizamento da forma, assim como um sutil deslizamento semântico, parecido com o que acontece no processo de aquisição do léxico por parte das crianças. Esse fenômeno, Heidi chama de *misparsing*: algo como “erro de avaliação sonora/semântica”. E as crianças utilizam essa técnica durante todo o processo: quando dizem [?áua] ao invés de [água], por exemplo. Estamos diante de uma nova entrada [?áua], ou diante de um deslizamento de forma? Esse processo de *misparsing* é tão poderoso que algumas formas ganham força suficiente para ser completamente substituídas por outras.

A noção *back-formations*, ainda segundo Heidi, é uma sub-espécie de *folk etymologie*. Nós, de uma outra forma, estamos assumindo aqui que a noção *back-formations* é o lado mentalista da formação de palavras e a *folk etymologie* é o dado sócio-histórico, capturado por processos descritivos de formação de palavras. Da conjugação desses dois processos, resulta uma entrada lexical totalmente nova entrando na língua. Essa conjugação ajuda, por exemplo, um aprendiz que, diante de uma palavra que contém uma sequência sonora nunca ouvida antes, pode atualizá-la semanticamente.

Uma palavra que não contenha um sentido buscado em *back-projection* (regras internalizadas semanticamente), na mente dos falantes de uma comunidade de fala, não será capaz de se sustentar em termos semânticos, por conseguinte, não terá forças suficientes para se manter como uma nova entrada. Em outros termos, uma nova entrada necessita ser parafraseada por entradas mais ou menos conhecidas pela comunidade – estar na mente dos falantes de uma comunidade – para que se sustente como entrada a permanecer numa lista.

Em razão disso, a criança-aprendiz lança mão de uma maquinaria mental para analisar uma nova entrada – e usá-la adequadamente, pelo menos em contextos razoáveis – e inventaria um sentido para a comunicação efetiva. De onde, portanto, se origina uma palavra como *invençoneiro*? Em termos de nossa argumentação mentalista, da capacidade para o *parsing*, herança mental oriunda da evolução da espécie.

Esse é um processo que acontece o tempo todo. Como as crianças são o tempo todo tocadas por palavras que nunca ouviram antes, e são hábeis em decifrá-las semanticamente, isso sinaliza uma evidência poderosa: a capacidade de *parsing* é genética. Suponhamos que uma palavra como [*triciclo*] seja posterior a uma outra chamada [*monociclo*]. A criação de um item como triciclo, objeto com três rodas que serve para locomoção, relaciona-se de uma certa forma com o item monociclo. O item triciclo foi aparentemente (hipoteticamente) motivado por uma formação retroativa: monociclo. O mesmo princípio de *parsing* e *back-projection* poderia ser utilizado para a criação de itens, tais como quadriciclo, octociclo etc.

Nesses termos, a mente junta itens aparentemente incompatíveis, mas mentalmente conciliáveis para as novas formações. Assim, ainda, dizemos que o léxico é gerativo pela capacidade composicional e pela maquinaria de *parsing* e *back-projection*.

Continuando, vamos a uma proposta também amplamente apresentada na literatura, qual seja, a noção de economia linguística. Mas vamos relacioná-la de uma maneira nova ao conceito de formação de novas entradas por redução e/ou ampliação semântica.

A redução e/ou ampliação semântica é uma outra poderosa estratégia para mergulharmos no processo de aquisição do léxico pela crianças. Geralmente há uma economia em termos formais, um item dado desliza semanticamente o suficiente para adquirir a condição de entrada nova. Como também uma forma perde determinados segmentos fônicos, mantendo ou não o mesmo valor em termos de conteúdo, passando, por vezes, a uma nova entrada lexical.

Nas palavras de Heidi, a palavra, reduzida em termos de segmentos fônicos – casos comuns como [*fotografia*] para [*foto*]; [*soropositivo*] para [*poz*] – é clipada (reduzida, diminuída) de tamanho. A clipagem de uma palavra é um processo de economia linguística

em termos formais e em termos semânticos. Os falantes descobrem que uma palavra reduzida de tamanho mantém o seu valor de conteúdo. Eles se dão conta de que não precisam usá-la de forma integral para recuperar o seu conteúdo semântico e identificar o seu conceito. Isso pode ser algo banal, até a gramática tradicional no capítulo sobre derivação pressupõe derivação por redução, mas, como o nosso foco aqui é mentalista, a clipagem, assim, é a estratégia fundamental no processo de construção lexical por parte das crianças em fase de aquisição do léxico.

Uma entrada como [ʔáua], por exemplo, é geralmente considerada uma nova entrada pois, digamos, é um deslizamento formal, uma fase intermediária, até ser cunhado o item [água] pelas crianças. Nesse estágio, fala-se em deslizamento por clipagem. Numa comunidade adulta, o deslizamento, em estágio mais avançado, é reconhecido, por falta de uma terminologia melhor, como derivação por redução. Nesse caso, podemos admitir, sim, que a redução cunha um novo item lexical. No quadro abaixo, resumimos as estratégias de aquisição de novas entradas pelas crianças.

ADQUIRINDO NOVAS ENTRADAS NA LÍNGUA		
FORMAÇÕES RETROATIVAS		
BACK-PROJECTION		
FOLK ETYMOLOGIES	Misparsing	Clipagem
DESLIZAMENTO SEMÂNTICO	Re-análise	Acoplagem

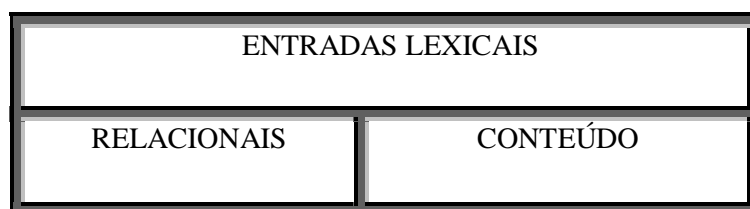
Quadro 6

Resumo: modulação e seleção, e como os falantes conseguem lexemas, foram objeto deste capítulo. Investiremos agora num estudo de caso, jogando com a entrada lexical na estrutura argumental.

## 6 A NOÇÃO DE ENTRADA LEXICAL NA ESTRUTURA ARGUMENTAL

FOCO: Entradas lexicais na estrutura argumental e no léxico, a noção de agente como a informação semântica em evidência.

Este capítulo, dividido em duas seções, tem como objetivo relacionar a noção de entrada lexical à estrutura argumental. Assume pressupostos da semântica lexical como base teórica e conceitual. O ponto 4.1 assume que todos os itens que compõem o léxico de uma língua são itens lexicais, e classifica-os em itens relacionais e itens de conteúdo.



Quadro 7

Tanto um (os relacionais) quanto outro (os de conteúdo) são chamados de itens lexicais por fazerem parte de um dicionário mental e público, que compõe o léxico, como já vimos em capítulos anteriores.

O objetivo específico deste ponto 6.1 é relacionar esses itens à definição canônica “estrutura argumental”, proposta no que na literatura semântica em linguística é discutido nos papéis temáticos, e estudado num projeto semântico chamado teoria temática, aludido acima. O ponto 6.2 tem como objetivo básico, dadas as definições de itens lexicais de conteúdo e itens lexicais relacionais e suas implicações de sentido com a estrutura argumental da sentença, problematizar as relações entre derivação morfológica de itens lexicais e estrutura argumental. Os dois pontos confluem para uma conclusão proposta neste capítulo: enquanto a flexão influencia no todo da organização sintática de uma sentença; a derivação, a partir de relações semânticas lexicais bem delineadas, desempenha papel fundamental na organização do sentido de uma sentença, apesar de influenciar semanticamente também na sentença, como veremos no próximo capítulo. Para antecipar as discussões, vejamos o seguinte exemplo:

61) O *joga-dor* atirou a bola para longe.

Em (61), há um morfema derivacional em itálico [-*dor*]. Ele faz parte do item lexical derivado [jogador]. Não é nosso propósito aqui delinear as ocorrências semânticas em termos de traços ou de leituras semânticas (Katz, 1973) dos morfemas da língua portuguesa. O leitor pode recorrer a Basílio (1980), Sandmann (1992) ou Rocha (1998). Estamos querendo dizer, basicamente, que um morfema da língua pode variar em leituras semânticas. É o caso, por exemplo, de [-eiro/a], nas ocorrências abaixo:

(62) macieira, mangueira, coqueiro, abacateiro, cajueiro

(63) tesoureiro, copeiro, barraqueira(i)

(64) peniqueiro, barraqueira(ii), fuleiro

É certo que em (62), em (63) e em (64), as leituras ou traços semânticos do morfema em foco [-eiro/a] variam. Delimitar uma leitura semântica para um morfema da língua é tarefa que pode ser designada pelas relações temáticas que os itens assumem numa sentença. Os itens necessitam “sair” do léxico (dicionário) da língua e se projetarem (KATZ e FODOR, 1969) numa sentença. Por exemplo, o item [barraqueira] é ambíguo entre, no mínimo, duas leituras. E essas duas leituras estão assinaladas com (i) em (63) e (ii) em (64).

Leitura I: o item pode tratar de uma atividade ou profissão de uma pessoa, como atesta a sentença abaixo:

(65) Joana é barraqueira, ela trabalha na feira, porém, somente aos sábados e aos domingos; nos outros dias, arma a sua barraca na praça.

Leitura II: ele pode ser avaliativo negativamente, por desnudar a personalidade nervosa de uma pessoa, como na sentença seguinte:

(66) Joana é a maior barraqueira, pois apronta confusão por onde quer que passe.



Os contextos sentenciais como articulados em (65) e (66) licenciam as leituras semânticas. Em (66), o traço avaliativo se projeta, fazendo com que o item “sofra” uma redução semântica. Tal redução se fixa ao contexto sentencial. Em (65), ocorre o mesmo. Há uma redução semântica do item, que, isoladamente, poderia se expandir em leituras (I) e (II), como proposto. Vamos assumir aqui que as relações temáticas são parcialmente essenciais para a projeção semântica dos itens derivados em estudo.

Por exemplo, em (61), o papel de agente, selecionado pelo verbo como seu argumento à esquerda, condiz com a leitura lexical agente para o item [jogador]. Nada mais plausível, e comumente aceito pelos falantes da língua, que o morfema [-dor] tem algo a ver com o papel de agente exigido pelo verbo. Mas esse morfema não perde esse traço semântico em qualquer lugar onde ele esteja disposto na sentença, e é a esse favor que vamos argumentar aqui.

É preciso, nestas alturas do que foi exposto, que invistamos nas seguintes perguntas, pois tomamos algumas coisas como tácitas até agora: por que é o verbo que seleciona o agente [jogador] em (61) e não o contrário? (Seleciona em termos relacionais, mas em termos de traços semânticos não é provável). O que significa uma regra de projeção semântica? Como itens tomados isoladamente numa língua, caso de [barraqueira], ambíguos em potencial, desambiguam-se numa sentença? Como é possível, por exemplo, que a entrada lexical [-eiro/a] que aparece em (65) não seja a mesma que aparece em (66)? Essas duas últimas questões têm a ver diretamente com regras de projeção semânticas, e com a estrutura argumental, e vamos discutir isso logo abaixo.

## 6.1 ITENS DE CONTEÚDO E ITENS RELACIONAIS

Vamos assumir, conforme uma literatura semântica exaustivamente proposta em linguística (CRUSE, 1992; HEIDI, 2005; BIDERMAN, 2001), dois tipos de itens lexicais que compõem o léxico de uma língua: a saber, *os itens de conteúdo* e *os itens relacionais*.

Os itens de conteúdo em um léxico são aqueles semanticamente independentes; fundamentalmente referenciais, e cuja categoria linguística assume determinadas posições numa sentença, tipicamente diferentes das posições assumidas pelos chamados itens relacionais.

Por itens de conteúdo, podemos listar algo nos termos: [surucucu, aparvalhado, luarão, galista, linguarudo, invencioneiro]. Já os itens de relação são aqueles semanticamente independentes, semi-independentes ou totalmente dependentes (de outros); relativamente, nunca ou quase nunca referenciais, cuja categoria linguística assume posições diferentes dos chamados itens de conteúdo numa sentença; e que, basicamente, servem para relacionar itens dentro de uma sentença ou dentro de um sintagma. Itens de relação cabem numa lista como esta: [é, está, com, para, quando, de, parece, andar, sair, comer, beber].

Esta proposta é próxima, de uma forma um pouco menos específica, à proposta de Cruse (1992), que define itens abertos e itens fechados numa língua. Aos abertos, ele atribui a noção conteúdo; aos fechados, ele atribui a noção gramática. É próxima pois elas não se recobrem totalmente. Mais adiante vamo-nos remeter à proposta de Cruse (*op.cit.*). Por enquanto, vamos investir no que delimitamos aqui como itens relacionais e itens conceituais.

Os itens relacionais são utilizados numa língua para conectar outros itens relacionais entre si, ou itens de conteúdo. Notem-se as sentenças que seguem:

(67) Ser é parecer.

(68) A galinha parece poedeira.

(69) A surucucu mordeu o menino.

É bom lembrar que não estão aqui em foco categorias gramaticais, e, sim, uma perspectiva semântica. Apesar de serem prototipicamente relacionais, por relacionar um item a outro, os itens [ser] e [parecer] estão sendo usados em (67) como conceituais; já em (68), o item [parecer] está sendo usado como relacional. Reitero que é bom esquecer momentaneamente a categoria gramatical, pois é evidente que em (67) [parecer] funciona como predicativo e em (68) [parecer] funciona como predicado. Em ambas as ocorrências, (67) e (68), contudo, não há mudança de traço ou de leitura semântica. Eles são projetados nas sentenças semanticamente com o mesmo significado. Em (69), o item [morder] é basicamente relacional ao estabelecer uma ponte entre os itens [surucucu] e [menino]. Mas o leitor deve notar que os itens [surucucu] e [menino] são prototipicamente itens de conteúdo. Eles não se prestam a fazer relações de tipo algum.

Quais as relações possíveis nas sentenças da língua portuguesa, se fôssemos isolar o verbo como o item relacional aqui em questão?

Elas podem ser à esquerda e à direita; podem ser somente à esquerda; somente à direita; ou às vezes um item prototipicamente relacional, dentro dessa nossa argumentação, assume a condição item de conteúdo. Vejamos isso com os seguintes dados:

(70) Serafina bebe vinho.

(71) Serafina deu a taça de vinho a Suane.

(72) Serafina sonha.

(73) É tarde.

(74) Chove.

Em (70), o item [bebe] relaciona um item que está à sua esquerda [Serafina] com um outro que está à sua direita [vinho]; em (71), o item [deu] relaciona um item que está à sua esquerda [Serafina] com outros dois que estão à sua direita [taça de vinho] e [Suane]; e em (72), o item [Serafina] se projeta na sentença relacionando-se na posição onde se encontra, à esquerda, com o item relacional [sonha]. Em (73), é o item à direita [tarde] que se junta ao item relacional [é]. Em (74), o item relacional [chove], como concebido aqui, é basicamente um item de conteúdo pois relaciona nada a nada. Ao seu lado estão dois vazios: um à esquerda e um outro à direita [ ] chove [ ].

Em concepções gerativas, argumenta-se que a flexão seria a categoria funcional que, na estrutura profunda, estaria à esquerda do item [chove], portanto relacionando-se com ele. Mas o leitor deve seguir a recomendação que propomos logo antes, e esquecer momentaneamente categorias sintáticas, concretas ou abstratas aqui. Para isso, montemos as seguintes ocorrências, para defendermos a condição-conteúdo ao item relacional [chove]. Em sentenças não-neutras<sup>23</sup>, tais como:

(75) Chove.

(76) Chuva.

Os itens em (75) e (76) se projetam sintática e morfologicamente de maneiras diferentes nas sentenças. Admitimos categoria flexional-funcional operando para o item em (75) [chove], e teremos que admiti-la também para o item em (76) [chuva]. Ou seja, se há flexão operando tempo (é inadmissível pensar na flexão-pessoa para o item (75) [chove], pois isso seria semanticamente uma anomalia: “ninguém” chove), há flexão para também com o item (76) [chuva], flexão, nesse caso, de número, mas se estivéssemos pondo em relevância projeções morfossintáticas.

---

<sup>23</sup> Não-neutra trata-se de uma sentença dentro de um contexto. Por exemplo, alguém olha para fora do apartamento e pode dizer tanto a sentença (75) Chove; quanto a sentença (76) Chuva. Em ambas as ocorrências, adequadas para esse mesmo contexto, os itens [chove] e [chuva] não se relacionam com outros itens, portanto, admitimos aqui que se trata de, tanto um quanto o outro, itens de conteúdo.

Tanto um item quanto o outro, contudo, não estabelece relação (pelo menos semântica) nem à esquerda nem à direita com nenhum outro item, e semanticamente eles podem ser usados em contextos similares.

Admitir itens relacionais e itens conceituais vai-nos servir ao seguinte propósito: os itens relacionais sentenciais (basicamente os verbos) selecionam itens conceituais (ou não) à sua esquerda e à sua direita. Estamos admitindo que alguns itens selecionam outros, e que os sentenciais (os verbos) são os selecionadores, e, portanto, desempenham papel central na estrutura semântico-argumental da sentença.

Vamos investir na argumentação do que acabamos de afirmar (os verbos são os selecionadores), e responder à pergunta posta logo acima e retomada aqui: por que é o verbo que seleciona o agente [jogador] em (61) e não o contrário? Vamo-nos basear nas argumentações de Chafe (1979) para a língua inglesa para responder a essa pergunta.

Toda sentença relevante na língua portuguesa, relevante em termos sintáticos, em termos semânticos, em termos discursivos e em termos pragmáticos para os falantes do português – sobretudo em termos linguísticos para fins meta-pedagógicos – é organizada em torno de um verbo. Muitas vezes esse elemento predicativo é circundado por elementos nominais. As ocorrências (70) e (71) são exemplos típicos. Vemos elementos relacionais-predicativos [bebe] e [deu] circundados por elementos nominais [Serafina], [vinho], [taça de vinho] e [Suane].

Chafe (*op. cit.*, p.96) divide “o universo conceptual humano”, em termos de organização de uma sentença, em duas áreas: uma área é a do verbo, que, ele continua, engloba estado (condições, qualidades) e eventos. A outra área é a dos nomes, que engloba coisas (objetos físicos e/ou abstrações coisificadas). Ele afirma categoricamente que uma área é central (a área do verbo) e a outra área é periférica (a área do nome). Discordamos disso, contudo, pois, em termos de uma semântica lexical, isso é pouco provável. Não existem hierarquias lexicais em nossas mentes. Em termos semânticos, todos os itens se equivalem, não existem itens-núcleos nem itens-periféricos em termos lexicais.

Uma checada rápida nos dicionários da língua portuguesa atesta que ali se encontram em verbetes coisas como verbos, advérbios, nomes, onomatopeias, itens como “a” etc. A única hierarquia é alfabética. Em termos de um dicionário mental, nem isso.

O interessante da argumentação de Chafe (*op. cit.*, p.96-97), no entanto, é que ele vai propor evidências para a centralidade do verbo para o que aqui nos interessa: a estrutura argumental das sentenças e a seleção dos itens dentro dessa estrutura, e, assim, o verbo é o item que seleciona os itens à sua esquerda e à sua direita.

Vamos remontar os argumentos desse linguista americano em três pontos defendidos por ele. O quarto ponto, nós deduzimos da proposta que apresentamos aqui:

A sentença que linguisticamente interessa contém necessariamente um verbo.

a)“Eloquções que semanticamente não têm verbo, como, talvez, Oh, Ah, parece ser melhor que se considerem como relíquias do tipo pré-humano (...) onde a regra era a simbolização direta de mensagens unitárias.”

Ele argumenta que isso – de uma sentença que interessa conter ao menos um verbo – pode ser um princípio de organização universal de uma sentença.

b)“A natureza do verbo determina como vai ser o restante da oração.”

Isso pressupõe uma gama de relações semânticas e categoriais. Sobretudo quais nomes ou constituintes o acompanharão, quais as relações desses nomes ou constituintes com ele, e, para o que nos toca mais de perto aqui, como esses nomes ou constituintes serão especificados semanticamente pelo verbo.

c) Os nomes não tendem a mudar a semântica dos verbos.

Tomem-se como exemplo, as sentenças que seguem:

(76) Chove muito em Salvador.

(77) Choveu papel picado no 2 de Julho em Salvador.

Se fôssemos parafrasear (76), diríamos que alguma coisa cai do céu, e o que cai do céu é a chuva. Se fôssemos parafrasear (77), diríamos que alguma coisa cai do céu, e o que cai do céu é papel picado. As sentenças seguintes, também, podem continuar nos ajudando a problematizar esse ponto:

(78) Serafina riu na aula.

(79) A parede riu de todos nós.

Se fôssemos parafrasear (78), diríamos que o item lexical à esquerda do verbo dessa sentença é animado, e praticou uma ação: riu na aula. Se fizéssemos o mesmo com (79), diríamos que o item lexical à esquerda do verbo dessa sentença é animado e praticou uma ação: riu de todos nós. Quem define o que é animado ou não numa língua não é um biólogo, e, sim, os falantes. Nesses termos, é necessário que atribuamos o traço [+animado] aos itens em questão para fazê-los se adequar à semântica do verbo que emerge nessas sentenças. E isso nos leva a crer que quem faz isso é o verbo, pois o traço [+animado] se encaixa na leitura semântica de quem pratica a ação [rir], e que necessariamente deve também se encaixar com a leitura semântica inerente ao item verbal relacional em questão.

O verbo é um item relacional: é ele quem indica semanticamente o quê se relaciona a quê.

Retomemos a sentença (69):

(69) A surucucu mordeu o menino.

Seria no mínimo estranho que os itens conceituais [surucucu] e [menino] mudassem de posição nessa sentença, tendo o item [mordeu] relacionando-os. É evidente que isso é categorial e semanticamente possível. Mas pouco provável, e uma sentença, digamos, pouco produtiva na língua. Temos certa desconfiança de que seria inusitada uma manchete de jornal com a seguinte estampa:

(80) Menino mordeu a surucucu.

Essa manchete exigiria determinados esforços interpretativos do leitor, não acionados se a estampa fosse a sentença (69). E percebamos que os itens conceituais [menino] e [surucucu] guardam traços parecidos compatíveis com a semântica exigida pelo item relacional [mordeu], ou seja, tanto [menino] quanto [surucucu] podem morder. Diríamos que o item-relacional [mordeu] comande as posições que os itens conceituais devem assumir nessa sentença. Assim, [mordeu] exige que o item [surucucu] se acomode à sua esquerda, e o item [menino] se acomode à sua direita.

Nesses termos, o verbo é o item relacional, digamos, por excelência, comandando os lugares semanticamente a ser preenchidos pelos itens de conteúdo numa sentença. O leitor deve estar atento ao fato de que estamos nos referindo a sentenças, e o verbo é esse item relacional por excelência das sentenças de uma língua, afinal, é a estrutura argumental em foco por aqui. Há outros itens de outra natureza que relacionam sintagmas, o que foge totalmente ao escopo de nosso trabalho neste capítulo.

Vamos trocar o item relacional [mordeu] na sentença acima, e por no lugar um outro item de mesma natureza. Isso acontece nas sentenças abaixo (81) e (82):

(81) A surucucu viu o menino.

(82) O menino viu a surucucu.

A estranheza semântica da sentença em (80) parece não ser sentida com tanta intensidade nas sentenças em (81) e em (82). Isso acontece já que os itens relacionais foram mudados. Em (80), o leitor atesta que a ocorrência era com o item [mordeu]. Em (81) e em (82), as ocorrências foram construídas com o item relacional [viu].

Estamos verificando, portanto, que a semântica do verbo nessas duas últimas sentenças faz a diferença nas leituras mais ou menos estranhas semanticamente em (81) e em (82), e totalmente estranha em (80). Ponto positivo para o argumento de Chafe, quando propõe que *os nomes não tendem a mudar a semântica dos verbos*; enquanto os verbos

tendem a mudar significativamente a semântica dos nomes, ou seja, as posições dos itens conceptuais se modificam em (69) e em (80), tornando a sentença em (80) estranha, ou totalmente estranha – quase anômala semanticamente. O que não acontece em relação a (81) e a (82), em que os nomes trocaram de lugar, mas como o item relacional é de outra natureza – é [viu], ao invés de [mordeu] – a estranheza das ocorrências é quase nula.

Conclusão provisória 4: os verbos, do ponto de vista estrutural e relacional, estão no comando semântico-relacional.

Os itens de conteúdo são essencialmente chamados de argumentos, um termo emprestado da lógica formal que foi introduzido nos anos sessenta pelos linguistas da semântica dos casos.

Os itens relacionais são chamados de predicadores, por cumprirem funções na dinâmica com os argumentos – essencialmente os verbos é que cumprem esse papel.

Como vimos acima, o número de argumentos nos arranjos entre os itens conceituais e os itens relacionais são os demonstrados na sequência de sentenças em (72), (73) e (74). A estrutura que as relações impõem aos conceitos é chamada de estrutura argumental.

### 6.1.1 Estrutura argumental

Há alguns tipos de estruturas argumentais que os itens relacionais podem impor. Definidos os conceitos no ponto 4.1, vamos então investir em alguns tipos dessas estruturas nos parágrafos que seguem.

O objetivo aqui não pode ser perdido de vista: às vezes um item precisa receber em relação o traço [+agente], é o caso do item [Serafina] em (83):

(83) Serafina viu o pedreiro

e, às vezes, esse traço é recebido em relação, é o caso de [jogador] em (61), assim como já vem marcado no léxico, é o caso também de [jogador] em (61), pois o sufixo [-dor] deixa esse traço semanticamente explícito, oriundo do léxico.

Não vamos aqui expor exaustivamente todos os tipos, pelo menos todos aqueles já consagrados na literatura sobre esse tema. Nada além do que aquelas estruturas já amplamente propostas em trabalhos de Jackendoff (1978), Filmore (1972), Franchi e Cançado (1996).

As relações tendem a impor alguns tipos de papéis a seus argumentos, o que significa dizer que as relações exigem argumentos com certas propriedades semânticas. Isso se realiza

somente parcialmente, e só é delimitado para as relações entre aqueles dois tipos de itens delineados no ponto 6.1, e nas sentenças.

O tipo mais emblemático de leitura semântica na estrutura argumental que uma relação pode ter é aquele tipo de argumento que faz algo a um outro argumento – o papel temático de agente é esse tipo. Exemplos de itens predicadores tipicamente agentivos são os que aparecem nas sentenças (61) [atirar]; (69) [morder]; (70) [beber]; (71) [dar]; e [arrasta] em (84), abaixo:

(84) O pescador arrasta a jangada.

O argumento que realiza a ação é o [agente]; em algumas propostas, como a de Franchi e Cançado (op.cit.), há ainda a noção [causativo], se se refere a seres coisificados tomados como animados.

Já o argumento sobre o qual recaem as consequências da ação, ou que tem a ação feita é o [tema], também reconhecido como [paciente].

As posições dos itens sobre os quais recaem propriedades argumentais, quando as sentenças estão na voz ativa, são típicas:

- 1) à esquerda do verbo, e categorizado como sujeito, é o [agente/causativo]; e,
- 2) à direita do verbo, e categorizado como o objeto, é o [tema/paciente].

Quando na voz passiva, essas posições se alternam, e quem está à esquerda recebe a propriedade [tema/paciente] e à direita, [agente/causativo]. Há casos em que, na voz ativa, o argumento [tema/paciente] pode também constar ao lado esquerdo do verbo.

É necessário, assim, um levantamento exaustivo das redes temáticas dos verbos para delimitar essas posições, pois cada verbo da língua portuguesa vai exigir leituras semânticas próprias na delimitação dessas suas redes temáticas. Isso é uma outra coisa que foge totalmente ao escopo do nosso trabalho aqui; contudo, mais adiante, pulverizamos um exemplo ou outro sobre a semântica de alguns verbos, e suas intervenções nas posições argumentais.

Quando propomos delimitar redes temáticas a que os verbos são submetidos em português, queremos dizer, basicamente:

- a) há um número de papéis (sentidos) que o predicado (o verbo) terá que atribuir;
- b) é necessário, portanto, estipular quais são esses papéis;
- c) como também dizer qual é o número desses papéis, a partir das relações semânticas que esses itens estabelecem nas sentenças, que pode variar de sentença a sentença.



Para não deixar as coisas talvez tão parciais assim aqui e agora, vamos jogar com a semântica do verbo [pegar] e algumas de suas demandas argumentais. Atenção às sentenças que seguem:

(85) Marivone pegou um táxi.

(86) Marivone pegou uma gripe.

(87) Marivone pegou o filho no colo.

(88) Marivone pegou a xícara sem cabo.

(89) Marivone pegou no batente.

Dadas as mudanças semânticas dos verbos nas ocorrências em acima, mudanças semânticas sutis, queremos crer aqui – de fato, não houve uma mudança semântica substancial – pois [pegar] em todas as sentenças do conjunto acima guarda alguns traços semânticos estáveis entre si, levantamos algumas assertivas, como também, algumas questões.

As assertivas:

a)em (88), o item conceitual [Marivone] parece exercer tipicamente o papel temático de [agente]: realiza uma ação, faz alguma coisa acontecer, é potência; e

b)em (86), o item conceitual [Marivone] parece experienciar algo, parece sofrer algo, parece ter sido vítima de algo; paradigmaticamente, o [agente] aí é um vírus.

As questões:

a)em (85), qual seria a leitura semântica argumental que poderíamos inferir para o item conceitual [Marivone]? Seria [agente]? Se assim o fosse, que tipo de ação esse argumento exerceu?

b)em (89), como poderíamos atribuir leitura argumental ao item conceitual [Marivone] aí, já que, parece, o verbo não pode ser tomado isoladamente na atribuição dessa leitura? Será que poderíamos assumir que o item à direita do verbo [no batente] não estaria sendo tragado pelo verbo, formando um conjunto fechado?

Essas são questões típicas de um estudo em grade temática dos verbos, e não é nosso propósito problematizá-las ostensivamente aqui. Elas, contudo, conseguem colocar na berlinda conceitual a disposição de Chafe (*op.cit.*) em admitir o verbo como pertencendo à área “central” de uma sentença bem construída numa língua.

Assim, queremos crer também que, o que esse autor diz, repetido aqui, qual seja, *a natureza do verbo determina como vai ser o restante da oração*, deve ser tomado com cautela, pois a sentença em (89) prova que é a natureza do verbo como também a natureza do

que vem depois do verbo influentes semanticamente para determinar *como vai ser o restante da oração*.

Notamos em parágrafos anteriores que uma relação pode envolver somente um argumento. Relações que envolvem apenas um argumento são classificadas amplamente de duas maneiras: uma em que o argumento único está no controle do que acontece, e uma outra em que o argumento único não está no controle do que acontece. Aqui, fazemos uma ressalva: temos uma aparente contradição com o que Chafe (*op.cit.*) propõe para os verbos. Lembre-se de que esse autor propõe, dentre outras coisas, que os verbos estão no comando das relações do ponto de vista semântico – e atentar para o argumento *o verbo é um item relacional: é ele quem indica semanticamente o quê se relaciona a quê*.

Como queremos, todavia, por em relevância pressupostos de uma semântica lexical, propomos que os verbos relacionam argumentos numa sentença, mas os itens conceituais relacionados pelos verbos também são relevantes para a semântica de uma sentença completa (ver logo abaixo o ponto 6.2, em que os itens derivados [+agente] controlam seu próprio traço semântico [agente]).

a) Casos em que o argumento único está no controle da ação expressa pelo verbo

Atentemos para as seguintes ocorrências:

(90) Serafina riu.

(91) Desesperadamente pela pradaria, o alazão galopa.

(92) Nara Costa canta no sábado.

(93) Suada, Suane corre para a aula.

O argumento único das sentenças em acima está no controle da ação. Temos aqui um jogo duplo em termos semânticos: o verbo exige uma propriedade [agente] para o seu argumento à esquerda (lembre-se do proposto por Chafe (*op.cit.*): *os nomes não tendem a mudar a semântica dos verbos*), e o seu argumento à esquerda assume o traço [agente] proposto pelo verbo para que a relação cumpra a demanda semântica da sentença: alguém/algo faz algo.

b) Casos em que o argumento único não está no controle da ação (conteúdo) expressa pelo verbo

Atentemos agora para as ocorrências seguintes:

(94) A planta cresce.

(95) A neve cai.

(96) O sol brilha.

Aqui temos exemplos típicos da semântica dos verbos proposta por Chafe (*op.cit.*): os argumentos não estão sob o controle da semântica verbal. Aqui, sim, os verbos controlam tanto a relação quanto o que seus argumentos estão fazendo, sentindo ou tendo a consequência de algo. Mas às vezes é difícil concordar com esse autor, pois demarcar as fronteiras de quem comanda quem do ponto de vista lexical não é tão previsível assim.

Em (94), (95) e (96), os argumentos são consagrados na literatura como o [tema] do conteúdo expresso pelo verbo. Atenção ao conjunto de sentenças que segue:

(97) A carta chegou.

(98) Suane dorme.

Quando os argumentos preenchem a propriedade semântica exigida pelo verbo, ou seja, em (97) [carta] está de alguma maneira em movimento, confundindo-se com um traço [animado], e em (98) [Suane] leva explicitamente o traço [animado], aqui, sim, temos casos em que os argumentos de Chafe (*op.cit.*) são aproveitados à risca, ou seja, quando ele afirma que *uma área é central (a área do verbo) e a outra área é periférica (a área do nome)*.

Nesses termos, os nomes, de fato, não modificam a semântica dos verbos, pelo contrário, os nomes sofrem as demandas semânticas exigidas pelo verbo.

Algo no mínimo problemático de se pensar quando temos, de um lado, o nome [cavalo], e, de outro, o verbo [galopar], dois itens lexicalmente mais próximos semanticamente que os itens [carta] e [chegar], em (97) e [Suane] e [dormir] em (98). Digamos que há uma “fricção” léxico-semântica de algum tipo entre [cavalo] e [galopar] que não encontramos entre os itens em (97) e (98).

c) Relações argumentais com dois argumentos e nenhum deles é o agente

Há relações argumentais onde nenhum dos argumentos é o agente. Em sentenças como as listadas abaixo:

(99) Serafina sabe crochê.

(100) Suane ama Mary.

Os itens-argumentos [Serafina] em (99) e [Suane] em (100) não estão, de fato, fazendo nada: não estão agindo como “agentes” costumam agir, nem tampouco os outros itens argumentos [crochê] em (99) e [Mary] em (100) não estão sofrendo consequências de ações dos argumentos à esquerda das sentenças onde aparecem, por não haver nem ações expressas pelos verbos dessas sentenças. Nesses casos, as relações parecem expressar sentimentos e

atitudes por parte de um argumento em relação a outro. Aqui, o argumento à esquerda é chamado de [experienciador], e o argumento à direita é chamado de [tema].

O que há de interesse nessas relações argumentais com experienciadores é que elas também permitem que seus argumentos denominados de [tema] possam expressar uma relação entre um experienciador e uma outra proposição. Assumimos como proposição aqui uma outra sentença ou um pedaço de uma sentença sobre o estado das emoções, do conhecimento ou dos sentimentos vividos pelo experienciador. As sentenças abaixo podem ilustrar melhor o que estamos afirmando:

(101) Serafina sabe que Suane está na aula.

(102) Serafina quer que Suane esteja na aula.

Ao invés de um [tema] em (101) e (102), temos uma proposição explicitando emoções, sentimentos ou conhecimentos dos argumentos à esquerda dos verbos dessas sentenças. Isso, contudo, não funciona quando os verbos prototipicamente selecionadores de [agente] como seu argumento à esquerda estão em cena. Veja, por exemplo, a anomalia de ocorrências como as que seguem abaixo:

(103) \*Serafina atirou que Suane está na aula.

(104) \*A surucucu mordeu que Suane está na aula.

(105)\* O pescador arrasta que Suane está na aula.

Assumimos, assim, como um axioma, e de forma sistemática, que itens relacionais que tomam experienciadores como argumentos à sua esquerda, ou que estejam inscritos no léxico como itens relacionais prototipicamente expressando sentimentos, conhecimentos ou emoções, além de relacionar experienciador a tema, relacionam também experienciador a uma proposição, como definida aqui em (101) e (102).

O mesmo não acontece, ou acontece parcialmente, com os itens relacionais que exigem um [agente] como seu argumento à esquerda, como atestam, negativamente, as sentenças (103), (104) e (05), e, positivamente, as ocorrências abaixo:

(106) Serafina falou uma palavra.

(107) Serafina falou que uma palavra não é o bastante.

Nessas ocorrências, um item relacional [falar], demandando um agente como um argumento à esquerda, relaciona esse argumento tanto a um tema [palavra] em (106), quanto a uma proposição [que uma palavra não é o bastante] em (107). Reiteramos, no entanto, que isso acontece de maneira parcial, ou seja, não com maioria significativa dos itens relacionais que demandam um agente como argumento à esquerda.

Com itens relacionais que demandam um experienciador como argumento à esquerda, esse fato da língua parece ser algo categórico e sistemático, no entanto, precisaríamos de mais testes com dados exaustivos da língua para sustentarmos essa posição.

d) Traço agente no léxico

Consideremos a sequência de itens abaixo:

(108) pasteleiro, carvoeiro, carpinteiro, serralheiro, marceneiro, marmoreiro, relojoeiro, padeiro, pipoqueiro, pedreiro

(109) pescador, controlador, armador, limpador, carregador, animador, pregador, curador, arranizador, predador

(110) ouvinte, cadeirante, fumante, agente, falante, escrevente, farsante, participante, ajudante, pedinte

Nas ocorrências acima, temos uma sequência de itens nominais inerentes ao léxico da língua portuguesa falada no Brasil, e todos de uma forma ou de outra são derivados e/ou nominalizados, deverbais ou não, portando os seguintes sufixos: [-eiro], para (108); [-dor], para (109); e [-nte] para (110). Vamos tomar como tácito aqui que em todas as ocorrências, esses sufixos portam o traço e/ou a leitura semântica lexical [+agente]. Vamos considerar agente aquele que realiza ações (tornam ações reais), pratica coisas, faz algo acontecer, é uma potência. Quando queremos tomar como tácito que todos esses sufixos aqui são de uma forma ou de outra [+agente], estamos propondo eliminar leituras semânticas, digamos, paralelas, como, por exemplo:

a) leituras avaliativas negativamente – é o caso de uma possível leitura avaliativa para o item [armador], no sentido de alguém que maquina artimanhas para prejudicar outrem; e

b) como também leituras que tomam objetos – é o caso específico do item [limpador] – como agentes em potencial, no sentido de coisa animada.

Fixada a leitura semântica como [+agente] para os itens em estudo, numa sentença, contudo, seu traço semântico não desaparece em qualquer lugar da sentença em que esse item estiver. Observem-se as sentenças que seguem:

(83) Serafina viu o padeiro.

(111) O padeiro viu Serafina.

(84) O pescador arrasta a jangada.

(112) A garota é cadeirante.

É difícil acreditar que o traço do item [padeiro], marcado no léxico como [+agente], inerente a esse item, desapareça, apesar de ser marcado como [+tema] para a sentença. O item [Serafina] é agente para a sentença em (83), mas [+tema] para a sentença em (111); já em (111) e em (84), vemos a confluência relacional da demanda verbal por uma leitura [+agente] *pari passu* com a leitura marcada [+agente] no léxico para os itens [padeiro] em (111) e [pescador] em (84). A sentença em (112) é emblemática do papel de [+agente] que o léxico pode instituir. O leitor percebe que o item em questão [cadeirante] nem faz parte de uma sentença em que exista um verbo demandando [+agente] como seu argumento, mas ninguém em sã consciência deixaria de fazer a leitura [+agente] para esse item. Ele é emblemático de uma luta antiga de pessoas com necessidades especiais em ser reconhecidas como produtivas, participantes da vida social.

Certa vez, nós nos referimos a uma pessoa numa cadeira de rodas, utilizando a seguinte assertiva: “Ela é tão bonita, mas, infelizmente, está presa a uma cadeira de rodas”; e logo fomos desautorizados por um interlocutor: “Você está totalmente equivocado, ela consegue ir e vir, deslocar-se, ir para outros lugares, enfim, ela é livre por causa de uma cadeira de rodas”. E logo cremos que um *cadeirante* faz coisas, realiza ações, portanto é potência e, por conseguinte, é agente.

Conclusão provisória 5: é preciso colocar o traço semântico [+agente] sob rasura quando é pressuposto como argumento de um verbo, pois há itens que, mesmo não sendo argumentos de um verbo que demanda [+agente], mesmo assim continuam sendo [+agente].

Há itens lexicais prototipicamente [+agente], é o caso de todos os itens em foco em (83), (111), (84) e (112) elencados acima, e itens que recebem o traço [+agente] em relação, é o caso dos itens [Serafina] em (83) e [garota] em (112).

## 6.2 REGRAS DE PROJEÇÃO SEMÂNTICAS E ENTRADAS LEXICAIS

“As pessoas reconhecem muito mais palavras do que as que empregam num determinado período de tempo.” Essa quase óbvia mas instigante assertiva de Pinker (1994, 182) leva-nos a discutir o que propomos logo acima no ponto 4.1 deste capítulo em forma de questões, aqui retomadas:

- a) O que significa uma regra de projeção semântica?
- b) Como itens tomados isoladamente numa língua, caso de [barraqueira], ambíguos em potencial, “desambiguam-se” numa sentença?
- c) Como é possível, por exemplo, que o [-eiro/a] que aparece em (65) não seja o mesmo que aparece em (66)?

Dados os limites da memória em armazenar centenas, milhares de entradas, como se houvesse a possibilidade de construir uma lista de itens interminável em nossas mentes – como se fosse uma lista telefônica que encaminharia o surgimento de um novo item, como se faz com os números – uma combinação incessante que se dirige ao infinito e que teria espaço constante para um novo item surgido –, vamos admitir que não lidamos simplesmente com entradas no nosso cotidiano uso de uma língua: nós lidamos com um léxico, que pode ser focado a partir de alguns critérios:

I) Há um ângulo público e social do léxico, dicionários tipo Aurélio e Houaiss são exemplos típicos desse tipo de domínio público, vimos isso em capítulos anteriores; e

II) se visualizássemos o léxico por um outro ângulo – se as lentes se virassem para dentro de nossas mentes – diríamos que em cada um de nós existe um léxico, mais ou menos talvez melhor, maior ou mais abrangente, mais restrito ou menor do que aquele que existe na comunidade de fala de que fazemos parte.

Seja como for que joguemos com o léxico que dominamos para construir sentenças, *as pessoas reconhecem muito mais palavras do que as que empregam num determinado período de tempo*, porque, para isso, elas atualizam *padrões*.

Vamos admitir aqui, já que a perspectiva deste capítulo, é, sobretudo, teoricamente em semântica lexical, que as regras atualizadas são regras de projeção semântica. Simplificadamente, podemos também admitir que os falantes “sacam” palavras de seu léxico social e mental e constroem suas sentenças. Por outro lado, os falantes também reconhecem palavras novas, atribuem significados a elas, pois eles atualizam regras de significação, ou seja, cada um dos falantes não precisa saber um significado de um item a todo momento, eles significam itens a todo momento pois eles também conseguem projetar significados. Por exemplo, suponhamos que ignoremos os seguintes itens do contexto abaixo:

(113) pastel-eiro, carvo-eiro, carpint-eiro

Do contexto (113) “não sabemos” o significado dos itens [pastel-], [carvo-] e [carpint-], mas sabemos o que significa o item [-eiro]. Agora, dado o contexto em (114), abaixo:

(114) tax-ista, contabil-ista, reform-ista

Suponhamos que saibamos o significado dos seguintes itens: [tax-], [contabil-] e [reform-]. Quando um item novo como [tax-eiro] surgir pela primeira vez à nossa frente, no mínimo podemos inferir seu significado, pois jogamos com regras de projeção semânticas, já que

temos internalizado o significado do item [-eiro] em (113), e o significado do item [tax-] em (114).

É por isso que não lidamos com o léxico como uma lista fixa e interminável, item a item. Atualizamos o significado das palavras jogando com regras de projeção semânticas. Qual a natureza dessas regras e quais os mecanismos de jogo que efetuamos com elas na montagem dos significados de que somos capazes cotidianamente? Para responder a essas duas questões, vamos recorrer ao trabalho dos linguistas americanos Katz e Fodor, “Estrutura de uma teoria semântica” (p.81-82). Nesse artigo, esses linguistas assumem quatro pontos característicos para conceber e delimitar tais regras de projeção, a saber:

a)“É preciso postular regras que projetem o conjunto infinito de sentenças de maneira a que espelhem o modo pelo qual os falantes compreendem sentenças novas.”

Concorre para a projeção do conjunto de sentenças uma formulação da língua que consiga conjugar uma caracterização gramatical e semântica da língua em questão, e aqui estamos desenvolvendo uma caracterização de tipo semântico.

b)“Ao encontrar uma sentença nova, o falante não está diante de elementos novos, mas diante de uma nova combinação de elementos conhecidos.”

Os linguistas investem o tempo todo em regras de projeção em relação a sentenças da língua, já que eles estão lidando com uma concepção semântica ao nível da sentença. Mas, como vimos, as regras de projeção atuam em todos os níveis da produção linguística, desde níveis mais inferiores – tipo a derivação – como vimos nos exemplos demonstrados pelos itens em (113) e em (114), em que o conhecimento de itens individuais ajuda o falante a inferir o significado de itens supostamente novos, como foi o caso de [tax-eiro] – a níveis mais avançados (é o caso de sintagmas, que veremos logo abaixo), e níveis superiores da produção (é o caso do axioma fundamental de uma língua, a saber, uma sentença).

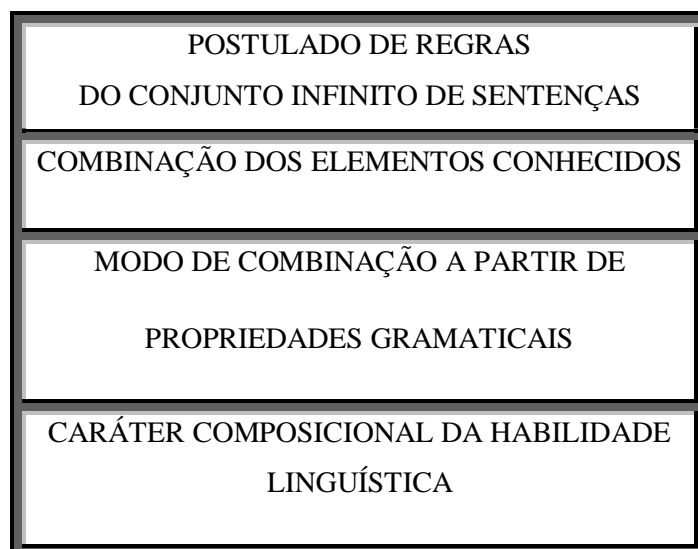
c)“Com base no seu domínio das propriedades gramaticais e dos significados dos morfemas da língua, as regras que o falante conhece permitem-lhe dominar o significado de uma nova sentença a partir do modo pelo qual as partes da sentença se combinam para formar um todo.”

Na aquisição da língua, os falantes mirins já conseguem ter plena consciência do domínio dessas propriedades gramaticais e dos seus significados, já que formulam a sua gramática ao receber os dados e articulá-los por regras em suas mentes. Digamos que eles recebem dois tipos de coisa: os dados, ou seja, o léxico de sua língua, e as regras, ou seja, como combinam esses dados, ao mesmo tempo.

d)Supõe-se que “um sistema de regras, para ser capaz de solucionar o problema da projeção, deva refletir o caráter composicional da habilidade linguística do falante”.



Esse caráter composicional da capacidade de produção de sentido pode ser visualizado a partir da forma como os itens se combinam em blocos, digamos, semânticos, até formar blocos maiores – estamos nos referindo aqui a constituintes semânticos, esboçados a partir de sintagmas de uma língua. No quadro abaixo, resumimos a delimitação de regras de projeção semântica.



Quadro 8

Os sintagmas abaixo ajudam a iluminar o que estamos querendo dizer. Vejam-se os seguintes sintagmas:

(115) \*[um cachorro falante]

(116) \*[uma pedra que canta]

(117) \*[o bebê honesto]

Ninguém em sã consciência levaria a sério semanticamente os sintagmas em (115), (116) e (117), pois todos em sã consciência estamos aptos a admitir que se trata de sintagmas anômalos semanticamente em língua portuguesa. Vamos acordar antes que tais sintagmas estejam fora de contextos poéticos e livres de contextos metafóricos.

Trata-se de anomalias, pois em (115) \*[um cachorro falante], a semântica de uma atividade como [falar], que deriva uma qualidade como [falante], não é compatível com as habilidades de um item que tenha leitura semântica como aquela que está presente no item [cachorro], ou seja, um cachorro “não fala”; já em (116) \*[uma pedra que canta], a proposição [que canta] é incompatível com um item não-animado como [uma pedra]; e, por fim, em (117) \*[o bebê honesto], de fato, esse sintagma “não funciona” semanticamente pois

o item [bebê] jamais poderia possuir uma qualidade como a expressa pelo item [honesto], perfeitamente possível para o sintagma [velho honesto].

Mas é possível salvar semanticamente os sintagmas em (115), em (116) e em (117). O que faz com que isso seja possível? Justamente o “caráter composicional da habilidade linguística do falante”. Vamos propor, então, salvar esses sintagmas, e o leitor fique atento que, ao salvá-los semanticamente, não estamos emitindo nenhuma sentença em contexto metafórico e/ou poético.

(118) Joana comprou um cachorro de pelúcia no Japão, funcionava a pilhas comuns, era [um cachorro falante].

(119) Encontrar [uma pedra que canta] seria a mesma coisa que encontrar um político honesto no Senado brasileiro.

(120) Cléber colou uma etiqueta em que estava escrito o nome honesto num cartaz onde estava o nome bebê, o cartaz ficou estranho, ficou assim: [bebê honesto].

A capacidade composicional de construção de sentenças também se observa em construção de itens em níveis mais inferiores da produção linguística, como a derivação e os sintagmas, e isso contribui para responder a duas questões que propusemos logo no início deste texto, retomadas no início deste item, e repetidas aqui, a saber:

*a) Como itens tomados isoladamente numa língua, caso de [barraqueira], ambíguos em potencial, “desambiguam-se” numa sentença?*

Esses itens desambiguam-se, pois, ao serem projetados numa sentença, o seu sentido individual já, em parte, “não lhes pertence mais”: o seu sentido é refeito, e, agora, surge uma nova forma de significação para eles – ou seja, da fricção do item [barraqueira] com os demais itens ao longo da sentença, como em (66), repetida aqui:

(66) Joana é a maior barraqueira, pois apronta confusão por onde quer que passe.

Vê-se claramente, esse item perde traços potencialmente direcionados para mais de uma leitura semântica – por isso sua ambiguidade – e é reduzido ao sentido compatível semanticamente para toda a sentença.

*b) Como é possível, por exemplo, que o [-eiro/a] que aparece em (65) não seja o mesmo que aparece em (66)?*

Ora, isso é possível, mais uma vez, dada a capacidade de concatenação composicional da produção linguística inerente aos falantes de uma língua. Nesse ponto, discordamos de Rocha (1998) quando ele afirma que um item como [-eiro/a] tem mais de um sentido. Um

item como esse tem mais de um sentido do ponto de vista de sua fricção semântica com os outros itens a que se acopla. Retomemos os exemplos em (62), em (63) e em (64), abaixo:

(62) macieira, mangueira, coqueiro, abacateiro, cajueiro

(63) tesoureiro, copeiro, barraqueira(i)

(64) peniqueiro, barraqueira(ii), fuleiro

Há um traço em comum para [-eiro/a] no conjunto de sentenças em (62), (63) e em (64): o traço comum é que esse item forma palavras derivadas na língua portuguesa – por falta de palavra melhor, vamos admitir que esse traço comum seja gramatical. Só que o seu sentido só pode ser deduzido em relação. Ou seja, é da fricção de [-eiro/a] com itens como [macie-], [mangue-], [coque-], [abacate-] e [caju-], que podemos atribuir o sentido de [-eiro/a] a algo como *item sufixal formador de palavras do tipo árvores*. Assim, é da fricção de [-eiro/a] com itens como [tesour-], [cop-], [1barraqu-], que podemos atribuir o sentido de [-eiro/a] a algo como *item sufixal formador de palavras do tipo profissão*. Por fim, é da fricção entre [-eiro/a] com itens como [penique-], [(ii)barraque-] e [fule-] que podemos atribuir o sentido de [-eiro/a] a algo como *item sufixal formador de palavras do tipo avaliativas*. Com uma ressalva aqui, um item como [barraqueiro(ii)] só assumiria o traço avaliativo em fricção com outros na projeção do todo de uma sentença.

Vamos imaginar também uma outra forma de pensar os itens lexicais de uma língua. No ponto 6.1, admitimos que esses itens podem ser semanticamente relacionais e de conteúdo, e numa série de argumentações demonstramos isso. Dissemos também que essa distinção em relacionais e de conteúdo é bem parecida com uma outra estabelecida pelo linguista americano Cruse (1992), ao separá-los entre itens lexicais e itens gramaticais, e admitimos que essas duas concepções – a saber, relacional/conteúdo e gramatical/lexical – são parecidas mas não se recobrem totalmente. É preciso, portanto, caracterizar mais explicitamente os itens gramaticais e os itens lexicais.

Isso é possível ao se observar o “comportamento” linguístico em termos de um conjunto de itens fechados (gramaticais) e um outro, um conjunto de itens abertos (lexicais). A que Cruse (*op.cit.*) define como, respectivamente, *closed set items* e *open set items*.

— *Closed set items*: são aqueles cuja classe faz parte dos itens que virtual e de fato são constantes durante a vida de um falante. Nesse conjunto, não há possibilidade de renovação dos itens, como também, numa sentença, há pouca ou nenhuma possibilidade de substituição. Observe a sentença abaixo:

(121) A garota o ama intensamente.

Compreendem esses *closed set items* afixos em negrito, como em: (garota-**a**, am-**a**, intensa-**mente**); assim como palavras como artigos, preposições e conjunções; a função linguística desses *closed set items* é sinalizar a organização gramatical da sentença. E essa organização gramatical é muito similar àquela já demonstrada pelo que ficou conhecido aqui como itens relacionais. Enquanto um organiza itens internamente (os gramaticais), o outro organiza sentenças (os relacionais).

— *Open set items*: são aqueles elementos que pertencem a uma classe que está sujeita a uma renovação constante de seus itens, assim, novos termos podem ser cunhados e outros podem cair em desuso.

Compreendem esses *open set items* raízes lexicais e palavras que portam o sentido principal dos vocábulos numa sentença, e em (121) são: garota, ama, intensamente. E esses itens do sistema aberto da língua são muito parecidos com os itens de conteúdo propostos neste capítulo.

Resumo: Os falantes projetam entradas lexicais de suas mentes a partir de noções intuitivas de itens relacionais e itens de conteúdo.

## 7 ENTRADA LEXICAL E DERIVAÇÃO

FOCO: Este capítulo se centra na delimitação do conceito de entrada lexical e suas relações com a noção de derivação, assim como estabelece o conceito de entrada lexical ao nível morfolexical, tendo sufixos agentivos como os dados para a delimitação de algo nos termos “entrada-sufixo”.

Neste capítulo, pretendemos o desenvolver um sistema descritivo que acreditamos ser fundamental para a discussão de aspectos semânticos da noção de entrada lexical na formação de palavras. Para que o sistema seja o mais orgânico e abrangente possível, alguns fatores serão levados em consideração:

- a) permitir que propriedades semânticas, que se espalham por variados níveis da língua, de átomos linguísticos, a palavras, recebam um tratamento em termos de produtividade lexical;
- b) propor a noção de entrada lexical e derivação conjugando pressupostos da semântica lexical com concepções mentalistas;
- c) observar o funcionamento da seleção semântica e da seleção categorial ao nível da noção de entrada lexical;
- d) estabelecer possíveis padrões semânticos que motivem a delimitação de entrada lexical.

Os dados, selecionados por intuição, tiveram que se submeter aos seguintes critérios:

- selecionaram-se somente sufixos [-eiro], [-dor], [-ista] e [-nte] formadores de nomes que recebem variados traços semânticos (aqui em relevância o traço +agente) e categorias gramaticais, notadamente a nominalização; e
- excluíram-se contextos de uso.

Vamos desdobrar as propostas de (a) a (d) em breves sub-capítulos, e observar no comportamento dos dados, padrões cognitivos de derivação. Essas propostas, à primeira vista tão amplas, estão muito próximas umas das outras. O tratamento lexical (a), por exemplo, se impõe dada a natureza do tipo de dado em questão; as noções de seleção semântica e categorial (c) e concepções mentalistas (b) se relacionam assim como os métodos (a composicionalidade) se relacionam com as teorias (as ciências da mente, ou da cognição); e, por fim, padrões semânticos (d) serão estabelecidos tendo essas propostas como base.

## 7.1 DERIVAÇÃO E LÉXICO



A formação de palavras, tema recorrente e intrigante nos estudos linguísticos, em princípio, pode ser amplamente definida como a atuação de um conjunto de unidades básicas motivadas por um conjunto de regras que combinam essas unidades em unidades complexas maiores, em camadas, para fazer surgir novas entradas em uma língua; ou ainda, um mecanismo sofisticado de expansão do léxico.

Vamos elencar algumas propostas de conceituação de formação de palavras, cada uma delas baseada em perspectiva teórica. Segundo Coseriu (1968, p.4), formação de palavras é “o conjunto de procedimentos mais ou menos regulares de que uma língua dispõe para formar lexemas secundários (derivados e compostos) a partir de lexemas que ela já possuía antes”.

Em consonância com propostas em morfologia gerativa (ARONOFF, *op.cit.*; SCALISE, 1986; COSERIU, *op.cit.*) admite regularidade precária e formação em cadeia, no sentido de que algo vem antes (lexemas que a língua já possuía antes) e algo vem depois (lexemas secundários), ao se acionarem Regras de Formação de Palavras (RFP's). Está em jogo a noção de primitividade. Isso precisa se fundamentar em dados de uma língua, pois, corre-se o risco de dogmatismos e pré-julgamentos, e mostra a fragilidade de pressupostos *ad hoc a priori*.

Preconiza-se uma RFP (Regra de Formação de palavra) (Rocha, 1998) do tipo  $[[X - ]_n \text{ ===== } X - \text{ eiro}]_n -$ , de *jardim* deriva-se *jardineiro*. A concepção de Coseriu (*op.cit.*) capta uma espécie de semi-regularidade: “(...) procedimentos mais ou menos regulares”. Pois, como vimos do conjunto de exemplos acima em 121, a regra capta a estrutura mas não é capaz de captar traços semânticos. Explicando melhor: *jardim* é um lugar, enquanto *jardineiro* é alguém que exerce uma função. Já *barraco* é um lugar, enquanto *barraqueira* é um termo avaliativo.

Já Cunha & Cintra (1987, p.83) assim se expressam:

Deixando de lado a viva controvérsia entre os lingüistas contemporâneos sobre a área de Formação de Palavras – se à morfologia, o seu domínio tradicional, se ao léxico ou à semântica, ou, mesmo se à sintaxe –, procuraremos tratar a matéria deste capítulo (...) com respaldo na seguinte conceituação...

Antes de enfocarmos a conceituação (um fragmento de Jean Dubois), propomos que a “viva controvérsia dos lingüistas contemporâneos” não seja deixada de lado, dado o prejuízo a um enfoque científico, fulcral para lidar com questões tão centrais como a Formação de Palavras. Vamos, então, à conceituação citada por eles:

Chama-se Formação de Palavras o conjunto de processos morfossintáticos que permite a criação de unidades novas com base

em morfemas lexicais. Utilizam-se, assim, para formar as palavras os afixos de derivação e os procedimentos de composição.

Buscamos na literatura o lado sintaxe da proposta de conjunto de processos morfossintáticos, e vimos que o mesmo tipo de regras que orientam a formação de um constituinte sintático, proposto por sintaticistas de correntes bem distintas (formalistas, estruturalistas e tradicionais) pouco se parece ao conjunto de regras que orientam a formação de uma entrada na língua.

O lado sintaxe da formação de palavras proposto por Cunha & Cintra (*op.cit.*), portanto, deve ser deixado de lado, pelo menos desde que a gramática das transformações da teoria padrão de Chomsky (1964) dos anos sessenta tornou-se obsoleta, e não passa, hoje em dia, de um fóssil linguístico.

O capítulo de Cunha & Cintra (*op.cit.*), que vai da pág. 83 até a pág. 115, não apresenta nenhuma proposta que evidencie o lado sintaxe dos processos de formação. Mas o que mais chama atenção nessa assertiva dos gramáticos Cunha & Cintra é a tentativa de focar a FP:

- a) como algo que cabe à morfologia ou ao léxico (vimos isso logo no primeiro capítulo de nosso trabalho aqui); e
- b) a base para a formação são os morfemas lexicais (observa-se que morfemas não-lexicais tipo afixo não gozam de relevância como os lexicais nos processos de formação).

Dos morfemas lexicais, vamos dizer que eles fazem parte de uma lista. Agora, o que fazer com a semântica é algo a ser incluído numa teoria de formação de palavras. Vamos tentar fazer isso mais adiante. Para tanto, vejamos como Cruse (2003, p.1) concebe a lista: essa “lista de unidades básicas constitui o léxico da língua; e regras combinatórias específicas constituem a gramática”. Vamos pressupor que as regras combinatórias de criação de novas palavras, constitutivas da gramática, têm motivação semântica.

As unidades lexicais básicas são duplamente observadas tanto por sua constituição (forma) quanto por seu conteúdo semântico. Forma e significado se juntam para que as regras sejam aplicadas com sucesso pelos falantes, que estão diante de duplo trabalho mental: eles assumem como bem sucedidas as formações que passam por um filtro tanto gramatical quanto semântico, tendo como base suas intuições e seu domínio subjacente da língua, ao mesmo tempo em que tendem a não aceitar as formações que não cumprem padrões estruturais de excelência das regras ou não cumprem bem a combinação entre seus significados.

Os falantes, à primeira vista, não se comunicam por entradas-afixos: eles se comunicam por entradas-palavras – um nível mais acima dos afixos, listadas no léxico. As entradas-palavras são tomadas por unidade apenas por um artifício didático ou pragmático,



um ponto de vista, ou uma lente de observação, mas é certo que uma entrada-palavra, dada a sua aparência multifacetada – o Capítulo II deste trabalho ilustrou que uma entrada-palavra recebe investidas de variados tipos para que esteja pronta para o uso – dificilmente não seja um objeto complexo. Complexidade exposta de forma inequívoca pelos padrões de derivação.

Vamos delinear a seguir que tipo de coisas as entradas-sufixos aparentam ser, e como os linguistas, em consonância com as intuições dos falantes, teorizam sobre elas. Tarefa muitas vezes lacunar, a noção de sufixo muda de língua para língua, como também de recorte teórico a recorte teórico, que por vezes põe em relevância sua forma ou classe, outras vezes seu conteúdo semântico, outras ainda seus aspectos em termos de homonímia e sinonímia, ou em termos de combinabilidade com outros sufixos, ou mais ainda em termos de sua produtividade.

Vamos deixar de lado alguns desses aspectos acima, e nos concentrar em torno de algo caro à FP: a noção de produtividade.

A noção de produtividade na formação de palavras, tema controverso, intuição antiga em linguística, foi introduzida na literatura de forma sistemática para uma teoria em morfologia gerativa a partir do trabalho pioneiro de Aronoff (1976). Em linhas bem gerais, delineamos a seguir seus marcos teóricos:

- É uma tentativa de delinear padrões ou regras que atuem de forma efetiva na formação de palavras, dado que embora persistentes e estranhas, muitas coisas são possíveis em morfologia, apesar de algumas serem mais possíveis de que outras;
- são propriedades específicas que tentam separar regras de formação de palavras com rendimento (em termos quantitativos), atuantes, daquelas não produtivas;
- é uma noção que nunca é inserida no estudo da sintaxe; dado que em sintaxe, claro, há certos tipos de operações que são imunes a questões de produtividade; por contraste, Regras de Formação de Palavras (RFP's) são sempre opcionais;
- segundo ainda Aronoff (1976, p.49), “Se quisermos comparar a produtividade de duas RFP's, podemos, simplesmente, listar palavras formadas por seus respectivos processos e encaixá-las (em suas RFP's). A lista mais longa será a lista da RFP mais produtiva”;
- a produtividade leva em conta restrições, bloqueios e escopo das regras, que funcionam em alguns casos, mas falham em outros.

A produtividade de uma regra de formação de palavras resulta de uma intersecção de um complexo de fatores. O mais proeminente deles atesta que a criação de novas entradas não é feita a partir de uma lista caótica e aleatória, há uma hierarquia baseada nas possibilidades e impossibilidades de junção de afixos a bases. O mais estranho e difícil trabalho dos linguistas é definir tais hierarquias, elencar as possibilidades reais e virtuais, e estabelecer critérios científicos aceitáveis para fixar uma regra como produtiva.

A par dessas dificuldades e estranhezas, há linguistas, no entanto, que rejeitam totalmente a noção de produtividade, associando-a a um delírio dos recortes de estudo da língua baseados numa perspectiva sincrônica, por confundir os interesses dos falantes aos interesses de um linguista.

Nessas análises, a competência do falante nativo é vista como um artifício epistemológico, ainda que útil, assentado em algo vago como intuição dos falantes, que se baseiam em julgamentos sobre o uso de sua língua frágeis o bastante para ser submetidos a um crivo científico.

Vamos apresentar abaixo argumentos nada desprezíveis que colocam em cheque a noção de produtividade de regras de formação de palavras, baseados no artigo “Problemas de morfologia e semântica histórica do sufixo –eiro”, de Viaro (2006):

- a fixação de regras de produtividade carece de fundamentação científica pois a previsão de formações futuras escapa a essas regras;
- as raízes da formação de palavras não podem escapar à historiografia das línguas, visto que a própria noção de derivação, fixada sob um ponto de vista estritamente sincrônico, é um paradoxo, já que se alguma entrada é derivada de uma outra, isso pressupõe assumir que algo vem antes e algo vem depois, ou seja, uma entrada nova é baseada na história dos itens, portanto, na história das línguas;
- acreditar na competência dos falantes nativos, nos seus julgamentos intuitivos, é como se fosse apostar no senso comum; e critérios científicos baseados em intuições não se firmam de forma consistente;
- afirmar que os falantes desconhecem a história de sua língua e, portanto, baseiam seu uso no aqui e no agora, é temerário visto que convivem numa sincronia vários níveis da língua, assimétricos historicamente;
- desprezar a língua como instituição, priorizando-a como sistema, é um prejuízo científico e desonestidade com a tradição;

- “(...) felizmente a assim chamada guerra entre Sincronia e Diacronia foi superada pelos estudos funcionalistas atuais. Hoje há maturidade suficiente para entender que a língua é sistema e instituição (RIO-TORTO 1998, 2004)” (Viaro, 2006, p.8);
- “(...) mesmo alguém que, hoje em dia, preconize ainda a primazia do elemento sincrônico, não poderia negar que, ao obter, por meio de comutação, um lexema [sapat-] e um morfema [-eiro], a partir de [sapateiro], chamar isso de derivação não deixa de ser paradoxal. Pois derivar pressupõe, na visão tradicional de onde vem o termo, que uma palavra provinha da outra ou que vinha antes da outra: [sapato] surgiu primeiro e daí veio [sapateiro]”;
- a diacronia é inerente à derivação, pois as regras de formação de palavras são projetadas percorrendo uma trajetória no tempo, nesses termos, são históricas.

Vamos propor a necessidade conceitual e metodológica da noção de produtividade, apesar de considerarmos as críticas de Viaro (*op.cit.*) pertinentes e conceitualmente necessárias. Vamos observar esses pontos de vista um a um:

- a) A fixação de regras de produtividade carece de fundamentação científica pois a previsão de formações futuras escapa a essas regras.

Ora, a aceitação de um novo item na língua é consequência de um domínio virtual das regras, pois só assim é possível que um item seja compartilhado por uma comunidade de fala sem que seja necessário recorrer a um dicionário todas as vezes em que um novo item é cunhado. Vide a controvérsia quando da criação do item [imexível], que jamais teve seu significado questionado, enquanto teve questionada a sua possibilidade de existência na língua.

- b) As raízes da formação de palavras não podem escapar à historiografia das línguas, visto que a própria noção de derivação, fixada sob um ponto de vista estritamente sincrônico, é um paradoxo, já que se alguma entrada é derivada de uma outra, isso pressupõe assumir que algo vem antes e algo vem depois, ou seja, uma entrada nova é baseada na história dos itens, portanto, na história das línguas.

Talvez seja difícil dizer qual dos itens vem antes, qual deles vem depois, dada a precariedade da metodologia utilizada, ou até mesmo se os dois vêm ao mesmo tempo. Mas a historiografia das línguas escapa sistematicamente ao simples falante, que vai fazendo as combinações por dominar regras.

- c) Acreditar na competência dos falantes nativos, nos seus julgamentos intuitivos, é como se fosse apostar no senso comum; e critérios científicos baseados em intuições não se firmam de forma consistente.

Nenhuma teoria em ciência aparece do nada, como se fosse algo vindo de um não-lugar. Uma teoria em ciência, particularmente em Linguística, capta as intuições do senso comum, sistematizando-as por intermédio de uma metodologia. Se a noção de competência do falante é lacunar, por outro lado ela não pode ser descartada.

- d) Afirmar que os falantes desconhecem a história de sua língua e, portanto, baseiam seu uso no aqui e no agora é temerário visto que convivem numa sincronia vários níveis da língua, assimétricos historicamente.

Bom, já afirmamos acima que os falantes não dominam a historiografia de sua língua de forma sistemática – como um cientista, e concordamos que os falantes a dominam de forma intuitiva, e é verdade que várias sincronias convivem ao mesmo tempo, de forma assimétrica. Isso reforça a noção de produtividade, visto que é possível num dado momento da língua observar quais morfemas foram mais ativos ou inativos.

- e) Desprezar a língua como instituição, priorizando-a como sistema, é um prejuízo científico e desonestidade com a tradição.

Estamos misturando aqui alhos com bugalhos: enquanto a língua como instituição tem um valor político, a língua como sistema tem um valor metodológico para determinados propósitos, e a formação de novas palavras é um deles.

- f) “Felizmente a assim chamada guerra entre Sincronia e Diacronia foi superada pelos estudos funcionalistas atuais. Hoje há maturidade suficiente para entender que a língua é sistema e instituição (RIO-TORTO 1998, 2004)” (Viaro, 2006).

Ora, servindo a seus propósitos, a noção de sistema e a noção de instituição não são vistas como incompatíveis, mas para a formação de palavras, a noção de sistema, com suas possibilidades de previsão e exclusão, é mais metodologicamente viável.

- g) “Mesmo alguém que, hoje em dia, preconize ainda a primazia do elemento sincrônico, não poderia negar que, ao obter, por meio de comutação, um lexema [sapat-] e um morfema [-eir-], a partir de sapateiro, chamar isso de derivação não deixa de ser paradoxal. Pois derivar pressupõe, na visão tradicional de onde vem o termo, que uma palavra provinha da outra ou que vinha antes da outra: sapato surgiu primeiro e daí veio sapateiro.”

A derivação não preconiza tão somente uma origem e um destino. Origem: sapat-. Destino: sapateiro. Capta, antes de mais nada, um processo, uma soma de itens. [sapat-] + [-eiro] = [sapateiro].

h) A diacronia é inerente à derivação, pois as regras de formação de palavras são projetadas percorrendo uma trajetória no tempo, nesses termos, são históricas.

É pedagógico, ou até mesmo viável ou possível separar diacronia e sincronia em termos metodológicos para fins de pesquisas em linguística? Essa é a pergunta que deve ser feita.

Nesses termos, acreditamos que a noção de produtividade se sustenta. E mais: sistematicamente ilumina aspectos da formação, como, por exemplo, a estranheza ou a familiaridade de itens como os que seguem:

(122) [invencioneiro] / [inventador] / [inventor]

Esses itens são possíveis. Um deles, em determinado contexto, funciona melhor que outro, ou soaria melhor que outro. Pode-se dizer que [invencioneiro] é mais “estranho” e “fantasioso”. O mesmo será considerado para quaisquer outros itens em quaisquer outras regras de formação de palavras (ROCHA, *op.cit.*; Basilio, *op.cit.*; RIO-TORTO, *op. cit.*, VILLALVA, *op.cit.*), produzindo palavras comuns e/ou potenciais. Os falantes, claramente, não estão usando simplesmente listas quando eles fazem essas combinações estruturais, formais ou semânticas, eles estão mostrando, antes de mais nada, evidências de ter acesso direto às suas intuições sobre determinados padrões de formação.

Essas intuições parecem expressar a provável competência de ter acesso a itens um a um, como a padrões de organização desses itens formando o seu acervo vocabular. Mas até agora estamos patinando no campo das intuições. É claro que nós não estamos interessados meramente na existência de intuições. Nem mesmo em como os falantes produzem isso (essa é uma tarefa muito mais ampla). Gostaríamos de explorar antes de mais nada, alguns fatores que influenciam os padrões de formação em termos de produtividade.

Se podemos aceitá-los, os julgamentos de valor dos falantes também concordam com a ligação entre produtividade e coerência. Pois os falantes vão dizer qual membro do par tal como invencioneiro/inventor soa algo menos agradável, qual deles poderia ter um sentido menos especial ou pragmático.

De uma outra forma, a correlação é perfeitamente razoável. O mais seguro é o que uma entrada do par quer dizer, ou o que há de convergente em termos de traços semânticos ou arranjo estrutural comuns a essas duas entradas.

Vamos propor tratá-las aqui como algo parecido com a divisão unidades lexicais e lexemas (vide capítulo II) com que os linguistas concebem suas análises sobre as entradas-palavras.

### 7.1.1 Entradas-sufixos como unidades lexicais e lexemas

Nas sentenças abaixo, aparecem ocorrências com o uso do sufixo [-eiro]:

(123)O sapat-eiro cobrou caro pelo conserto do sapato.

(124)O cande-eiro fica aceso durante toda noite quando faz frio na fazenda.

(125)Alguém chamou João de “peniqu-eiro”.

(126)Valéria mora no Terr-eiro de Jesus.

Em todas essas ocorrências, o sufixo [-eiro] apresenta conteúdo semântico diferente. Em (123), ele ajuda os falantes a interpretarem [sapateiro] como uma profissão; em (124), [candeeiro] é um objeto de um tipo (não é uma candeia, por exemplo), pois é do roçar de [-eiro] com [cande-] que o nome do objeto se explicita; em (125), está em jogo a avaliação negativa de uma função, em termos de classe social, a partir da leitura que a junção [peniqu-] e [-eiro] proporciona; já em (126), formação de nomes próprios também se beneficia da ação do sufixo [-eiro]. Essa é apenas uma pequena amostra da força semântica e categorial desse sufixo, e dada essa variação, o tratamento a essa entrada-sufixo se beneficiaria de uma divisão em termos de: sufixos como unidades lexicais e como sufixos-lexemas.

Podemos dizer, portanto, que a entrada-sufixo [-eiro] em (123) (124), (125) e (126) representa diferentes entradas-sufixos. Ela pode ser pensada como uma forma lexical em termos categoriais, em todas as ocorrências de (123) a (126) a categoria criada foi um substantivo e como uma entrada-sufixo (ele assume conteúdo semântico diferente em cada ocorrência de (123) a (126)).

### 7.1.2 Definindo um sufixo-lexema e um sufixo-unidade lexical

Para Cruise (*op. cit.*, p.37), um lexema “(...) são as unidades listadas num dicionário. Um dicionário fornece uma lista de lexemas de uma língua”. Um dicionário, para fins organizativos e descritivos, indexa palavras de uma certa forma. Acompanhe um exemplo de indexação de uma entrada (como fizemos no capítulo IV para a entrada [gros, grosse]), aqui associada à noção de sufixo como uma entrada. Por exemplo, no verbete:

*ENFANT* s.2 gên. Menino, menina; criança, infante, filho, descendente. (fig) Produto, resultado. LES \_\_\_\_S D'ADAM, os descendentes de Adão. PETITS- \_\_\_\_S, netos. \_\_\_\_ DE L'AMOUR, filho bastardo. CE N' EST PAS UM JEU D' \_\_\_\_, não é um negócio simples. \_\_\_\_ GÂTÉ, criança mimada. \_\_\_\_ TROUVÉ, enfeitado. (Dicionário de Francês: Francês/português, Português/Francês, 41ª ed., Editora Globo, S. Burtin-Vinholes, p.186).

A entrada (*enfant*) se apresenta como complexa (nesses termos, complexa aqui não se trata somente de palavra fruto de derivação ou composição, mas complexa em termos de investidas gramaticais, categoriais e semânticas) procurando-se mapear ou simular seus variados usos e ocorrências. Informa-se que ela tem conteúdo semântico, que pode ser usada em dois gêneros, o masculino e o feminino, sem que seja modificada a sua forma; sabemos que ela pode servir como base para a criação de novas unidades, mais complexas de que ela, pois estamos vendo a ocorrência: [PETITS-\_\_\_\_S] etc. A mente do falante também atua preconizando um sufixo como entrada: nesses termos entradas-sufixos podem ser estabelecidas em bases parecidas, que em termos bem gerais, vamos simular que sejam indexados da seguinte forma:

[-*EIRO*]. Sufixo que forma substantivos relacionados a profissão (carteiro), gentílicos (brasileiro), lugares (terreiro), árvore (abacateiro); (fig), avaliativo (peniqueiro).

Segundo Cruse, *op.cit.*, p. 64), em termos mais técnicos, “(...) um lexema é um conjunto relacional de significados associados a um conjunto de palavras formais”. Por que não pensar assim em termos, digamos, atomísticos? Por que não parafrasearmos um sufixo como se segue: uma entrada-sufixo é definida como um conjunto relacional de significados associados a uma entrada-palavra, disponível plenamente para o uso.

Por vezes, eles não estão claramente relacionados, é o caso de [-eiro] em sapateiro (aquele que faz/conserta sapatos), banqueiro (aquele que é dono do banco), terreiro (um lugar) e mamadeira (um objeto). Isso, entretanto, seria assinalado como diferentes entradas-sufixos. Em outros casos, uma relação pode ser facilmente intuída, é o caso de [-eiro] em padeiro, ferreiro e sapateiro, que pode ser parafraseada como X que faz. X que faz pão, X que faz sapatos e X que faz ferro. Esses casos serão considerados como pertencendo a uma mesma entrada-sufixo.

Os dicionários listam em seus verbetes as informações de variados matizes para uma entrada-palavra, como apontado acima, e quando a palavra é sufixada, essas informações levam em conta a ação do sufixo que acompanha a palavra derivada. De um *corpus* que cunhamos de cerca de 400 entradas com os sufixos em estudo neste trabalho, encontramos a

noção de +agente mais ou menos sendo estabelecida para todos os sufixos em estudo, nos seguintes termos:

[-dor]: sufixo agentivo que se acopla a bases verbais;

[-ista]: sufixo agentivo que se acopla a bases nominais;

[-eiro]: sufixo agentivo que se acopla a bases nominais; e

[-nte]: sufixo agentivo que se acopla tanto a bases nominais quanto a verbais.

Mesmo compartilhando a mesma forma, a leitura semântica de base é a leitura em primeiro plano para classificar esses sufixos como fazendo parte de entradas-sufixos diferentes. Aqui cabem questionamentos fundamentais para continuarmos com o raciocínio:

- 1) Quais propriedades gramaticais ou categoriais podemos inferir sobre as entradas-sufixos?
- 2) Quais propriedades semânticas podemos inferir sobre as entradas-sufixos?
- 3) Como podemos deduzir contextos semânticos holísticos para as entradas-sufixos?
- 4) Como podemos deduzir abordagens semântico localistas ou componenciais para as entradas-sufixos?

Essas questões, desdobradas nos quatro pontos acima, receberam respostas diferentes dados os interesses teóricos dos linguistas devotados à formação de palavras, em especial, à formação de palavras envolvendo entradas-sufixos. Vamos nos basear no artigo de Cruse “O léxico” (2001), traduzido por mim, para enfrentá-las.

Vimos acima que os dicionários apresentam entradas-palavras principais distintas separadas para lexemas distintos, mesmo quando eles compartilham as mesmas formas; mas grupos de significados relacionados são apresentados como fazendo parte de uma única entrada principal. Quando as entradas-palavras são relacionadas é porque se trata de um conjunto de formas que diferem somente em relação a afixos flexionais (tais como afixos indicativos de singular e plural para nomes, verbos e numerais; indicativos de gênero masculino e feminino para nomes; ou pessoa, passado, presente e futuro para os verbos).

Vamos retomar um exemplo de Cruse (*op. cit.*) para explicitar melhor o que foi dito no parágrafo acima – lexemas distintos relacionados a entradas-palavras. Cruse (*op.cit.*) propõe fazer isso tendo em vista uma entrada como [*posição*]. O capítulo II propõe uma distinção entre entrada lexical e lexema baseado num critério de gradação semântica: lexema é uma entrada vazia semanticamente – jazendo nas profundezas da mente dos falantes a que chamamos de limbo lexical. Nesses termos, num dicionário, encontraríamos sentidos diversos para a entrada [*posição*]. Em estado de dicionário – portanto ainda um lexema – não teríamos um sentido exato para a entrada [*posição*].



Depois de sofrer ataques semânticos e gramaticais, e ter-se deslocado do dicionário para a estrutura, ou seja, do dicionário (mental) para uma sentença, a entrada [posição] deixa de ser um lexema e passa a ser uma unidade lexical. Assim se expressa Cruse (*op.cit*, p.64): “É comum chamar distintos significados de *sentidos*. Mas o que é exatamente um complexo som-significado? Podemos chamá-lo de *unidades lexicais*”. Assim, reiteramos que estamos usando as expressões *lexema* e *unidades lexicais* respectivamente, como uma entrada em estado de dicionário, e uma entrada que sofreu perda e alçamento de traços semânticos e gramaticais assim que desembarcou numa sentença.

Quando propomos uma gradação semântica, estamos jogando ao mesmo tempo com o que um lexema e o que uma unidade lexical podem ser. Sentidos associados a uma entrada lexical sofrem uma gradação semântica: há uma certa distância semântica entre os sentidos que podemos associar a uma entrada como [banco], por exemplo: instituição financeira, móvel para decoração de uma praça, (banco) de areia, (banco) verbo etc. Retomando a entrada [posição] do exemplo de Cruse (*op.cit*), a distância não é tão grande: lugar, opinião, situação. Nesses termos, no caso de *banco*, estamos diante de lexemas distintos; já no caso de *posição*, consideramos que estamos diante de um mesmo lexema.

Encetamos essa argumentação para aplicá-la ao caso de entradas-sufixos, que se comportam de maneira parecida com a argumentação disponibilizada em termos de lexema e unidade lexical para as entradas-palavras: [-dor], [-ista], [-eiro] e [-nte] trata-se de entradas-sufixos diferentes; assim como [-eiro] avaliativo e [-eiro] lugar trata-se de entradas-sufixos diferentes também.

Fixados os pressupostos para se estabelecer os sufixos como entradas, vamos nos voltar para as questões levantadas acima.

1)Quais propriedades gramaticais ou categoriais podemos inferir sobre as entradas-sufixos?

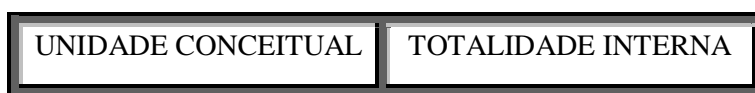
Em qualquer palavra em que as entradas-sufixos [-eiro], [-dor], [-nte] e [-ista] apareçam em língua portuguesa, essa palavra recebe o traço –verbo. São sufixos prototipicamente –verbo. Ainda, essas entradas têm uma posição fixa – ao final das entradas-palavras. Compete a ela receber as categorizações de gênero e número que incidem sobre toda a entrada-palavra. Se uma entrada-palavra estiver no gênero masculino, a marca deve ser encontrada na entrada-sufixo; o mesmo pode ser dito em relação ao número da entrada-palavra, cuja entrada-sufixo é o suporte para o número. As extensões prototípicas de uma

entrada-sufixo podem ser reconhecidas pelo fato de que elas são pedaços pequenos. E inserções que carregam informações como gênero e número respeitam suas fronteiras.

Partes das entradas-sufixos, portanto, não podem ser reordenadas (sapateiros; \*eirossapat; \*ssapateiro), não há liberdade de ordenação de entradas-sufixos para a língua portuguesa.

2) Quais propriedades semânticas podemos inferir sobre as entradas-sufixos?

Como vimos, há traços semânticos compartilhados pelas entradas-sufixos, objeto desta pesquisa, polissêmicos (relacionados) entre si, como também traços distantes semanticamente. Estamos convencidos de que os traços semânticos relacionados, assim como os traços semânticos divergentes ou distantes, comuns a uma mesma entrada-sufixo – é o caso do traço avaliativo e do traço [lugar], comuns a uma entrada-sufixo como [-eiro] dentre outros – podem ser sistematizados em termos de uma avaliação levando em consideração *a unidade conceitual e a totalidade interna*.



Quadro 9

Vamos delimitar a entrada-sufixo [-eiro] em termos de unidade conceitual e totalidade interna. É necessário considerar primeiramente que é prototípico dessa entrada a sua leitura categorial como uma entrada-sufixo formadora de nomes. Segundo, é prototípica dessa entrada uma leitura em termos de um feixe de traços semânticos relacionados entre si como também díspares entre si. Uma RFP acionada para criar uma palavra nova com esse sufixo é padrão: o produto final será sempre um nome. Nesses termos, sua totalidade interna nos dirá, sem medo de erro, que um substantivo será o produto final de uma RFP acionada para criar entradas-palavras com essa entrada.

O que chamamos aqui de *totalidade interna* pode ser mais facilmente ilustrado nos termos que seguem. Tome-se, por exemplo, as entradas [barraqueiro], [sapateiro], [açucareiro] e [bananeira]. Veja-se que há uma riqueza semântica em termos de traços semânticos: não fizemos nenhum levantamento, mas provavelmente, seja essa a entrada-sufixo mais rica em termos de traços semânticos em língua portuguesa, no sentido de variabilidade de opções de traços. A questão agora é saber quais desses traços são mais proeminentes, quais tenderão a entrar em desuso, quais são de maior ou menor produtividade.

Será possível que um desses traços ganhe a guerra semântica em relação a um outro e, por conseguinte, se torne o mais prototípico? Será possível que algum deles venha a ser

arcaizante? Fizemos um levantamento no Aurélio e vimos que o traço que aparece com mais frequência é aquele relacionado a profissão. Enquanto que o traço avaliativo pode estar em franca concorrência com a entrada [-inho].

Para entender algumas coisas desses processos, relacionados às questões acima, precisamos separar componentes dependentes e independentes de uma combinação semântica. O componente independente é aquele que rege as relações externas da combinação como um todo. Ou seja, todas as vezes em que uma entrada-sufixo como [-eiro] se acopla a uma base como [sapat-], a base é alçada na mente dos falantes como a informação que rege a combinação como um todo. Em outras palavras, fala-se de algo relacionado a [sapato] muito mais de que algo relacionado à entrada [-eiro].

A noção de componentes independentes nos permite estabelecer cadeias sucessivas de dependência semântica: não é simplesmente uma questão de se os elementos numa série são expressos com entradas dependentes ou independentes umas das outras, afinal de contas uma depende da outra para a significação total da entrada-palavra; noutros termos, não se trata de [sapat-] ou [-eiro], simplesmente, mas de algo presente na fricção entre essas duas entradas. Há uma incorporação de traços semânticos das duas entradas, como se elas se tragassem uma a outra.

[sapat-] + [-eiro] = *sapateiro*

[barraqu-] + [-eiro] = *barraqueiro*

[pipoqu-] + [-eiro] = *pipoqueiro*

[mangu-] + [-eira] = *mangueira*

[terr-] + [-eiro] = *terreiro*

[sorvet-] + [-eiro] = *sorveteiro*

Os elementos que constituem o significado total da entrada-palavra podem formar uma cadeia de sucessão contínua, mas, atenção, não é simplesmente o somatório de partes, é o entrelaçamento de partes das entradas para formar um todo. Pois, a entrada-palavra [barraqueira], simplesmente, não se trata de [barraqu-] + [-eira], mas de um deslizamento semântico da entrada [barraca] para algo que desemboca numa entrada-palavra avaliativa como [barraqueira]. Perfazendo, assim, uma *unidade conceitual*.

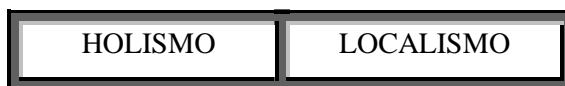
3) Como podemos deduzir contextos semânticos holísticos para as entradas-sufixos?

Segundo Cruse (*op.cit.*, 67):

De uma maneira bem ampla, prosseguir com a questão do estudo do significado depende do quadro que se estabelece para definir que tipo de

coisa o significado é. Algum domínio de questões maiores será útil como um background para discussões mais detalhadas que se seguirão. Podemos nos concentrar basicamente em dois pontos: holismo X localismo, e a relação entre significado lingüístico e conceitos.

Vamo-nos apropriar dessas concepções de Cruse para as palavras em Semântica Lexical e observar como podemos considerar a dicotomia holismo X localismo em termos de entradas-sufixos.



Quadro 10

O holismo é uma concepção que centra suas forças em definir o significado de uma entrada em termos relacionais, o que significa dizer que das relações de uma entrada com uma outra, no nosso caso, de uma entrada-sufixo [-eiro] com uma entrada como [sapat-] nasce o significado global da entrada palavra [sapateiro].

O localismo pressupõe que o significado de uma palavra entrada é auto-suficiente, e pode ser descrito independentemente dos significados das entradas com as quais se acopla.

Concepções holísticas sobre a linguagem centram suas abordagens em termos contextuais. Vamos observar como o contexto é tomado como base.

A dependência do contexto para determinar o sentido de uma entrada-sufixo considera a composicionalidade de uma entrada-palavra em que se encontra a entrada-sufixo. Noções de composicionalidade já foram abordadas em capítulo anterior. Nesses termos, a projeção de significado de uma entrada-sufixo não é exclusiva da entrada-sufixo de *per si*, mas, de um certo modo de combinação com os elementos que estão acoplados à entrada-sufixo numa entrada-palavra, ou seja, há uma dependência do meio ambiente lexical em que ela se encontra.

Pensemos, por exemplo, como um falante checa as entradas na mente antes do uso – noções de checagem estão mais bem sistematizadas em uma seção mais à frente. O fato de que um falante pode mobilizar uma entrada como [-eiro] para criar uma palavra nova, tal como [?disqueiro] – de [disqu-] + [-eiro] demonstra que ele compreende a noção de entrada em termos estruturais pois, como dissemos anteriormente, as entradas-sufixos não aparecem sozinhas na língua, mas são deduzidas da co-ocorrência com outras entradas-palavras, tais como [sapateiro] e [borracheiro], e prova que ele compreendeu o que, ou quais traços semânticos dessa entrada-sufixo podem ser mobilizados.

Em outros termos, o significado de [-eiro] (ou de qualquer uma outra entrada-palavra) não pode ser aprendido isoladamente. Lyons (1963) se fundamenta nesse *insight* básico, e

caracteriza o significado de uma entrada-palavra e, por extensão, de uma entrada-sufixo, em relação a sua posição numa rede de relações.

Vamos considerar o que isso significaria no caso de [-eiro]. Primeiro, [-eiro] pertence a um conjunto de entradas com as quais existe uma relação de exclusão, o que quer dizer que o que implica [-eiro] não implica [-dor], [-nte] ou [-ista] etc. Veja que todos eles compartilham o traço +agente. E todas essas entradas, por sua vez, pertencem a um conjunto cujos membros são mutuamente exclusivos, apesar de gozarem de um conjunto de propriedades gramaticais, tal como sufixos formadores de nomes.

Mas [-eiro] também mantém outros tipos de traços semânticos mutuamente exclusivos, tais como lugar, profissão, gentílico etc., ativando, portanto, outras redes semânticas. De fato, toda entrada-sufixo está ligada direta ou indiretamente, por meio de *links* específicos, como os traços semânticos acima demonstram. Nesses termos, para poder bem acionar o traço necessário para uso, o falante tem intuições ou conhece, e o significado de uma entrada-sufixo, nesses termos, não é inteiramente de todo compreendido até que todos esses *links* sejam conhecidos (embora, obviamente, alguns *links* sejam mais centrais de que outros).

Cruse (*op.cit*, p.72) ainda se expressa nestes termos em relação a concepções holísticas sobre o significado:

Uma versão alternativa da teoria contextual toma como origem o dístico de Wittgenstein: Não pergunte pelo significado, pergunte pelo uso. Isso é sugestivo, mas é também lacunarmente impreciso como base para a teoria do significado: o que se quer dizer, precisamente, com o termo uso? J. R. Firth (citado por Mackin (1978)) nos dá uma compreensão útil quando diz: As palavras podem ser conhecidas pelas companhias que as rodeiam. Essa linha de pensamento foi desenvolvida dentro da teoria holística de significado por W. Haas.

Resenhando as ideias de Haas, temos que uma entrada-sufixo disponível para uso e do domínio do falante foi semântica e integralmente normatizada, como está explícito na entrada [inventor], em contraste com [invençioneiro]. Ele então argumenta que se duas entradas diferem de sentido, esse fato inevitavelmente se refletirá na diferença de normalidade em um contexto ou outro. Por exemplo, há uma diferença bem sutil de significado entre *pé de banana* e *bananeira*, cujos contextos *pé de banana faz sombra* e *bananeira faz sombra* são capazes de explicitar: *pé de banana faz sombra* dá a impressão de unidade, enquanto *bananeira* dá a impressão de mais de um pé de banana.

Haas prosseguiu caracterizando o significado de uma entrada através de um perfil de normatibilidade por meio de todos os seus contextos de boa formação, reais ou potenciais:

nesses termos sinonímia plena de uma entrada-sufixo, do seu ponto de vista, possui o mesmo grau de normatibilidade em todos os seus contextos.

4) Como podemos deduzir abordagens semântico localistas ou componenciais para as entradas-sufixos?

Segundo Cruse (*op.cit.*, 75):

Um localista acredita que o significado de uma palavra é uma entidade auto-suficiente que pode ser em princípio finitamente descritível. Enquanto os holistas tendem a ver o significado de uma palavra como um conjunto de relações, tanto com outras palavras ou com contextos possíveis, um localista tipicamente dirá que essas relações são consequência do significado da palavra em si.

O naturalismo de Fodor (2004) é a faceta mais radical em termos de propostas localistas, que retrata o significado de uma entrada como uma montagem finita de partículas elementares de significado, e cada significado conta com algum aspecto do comportamento semântico do todo. Esses átomos semânticos (variavelmente conhecidos como *componentes semânticos*, *traços semânticos*, *marcadores semânticos*) são extraídos de um inventário finito, e nas versões mais fortes da teoria são psicologicamente reais (no sentido de que se sabemos o suficiente sobre o cérebro, estamos habilitados a identificar uma estrutura neuronal correspondente a algum traço), e universais (no sentido de que fazem parte da capacidade da linguagem de que cada ser humano é dotado ao nascer).

Adaptando os exemplos elencados de Cruse (*op.cit.*) para as palavras, levando em consideração o recorte conceitual localista, podemos apresentar algo nos termos:

[-eiro] = [nominalizador] [masculino] [avaliativo/lugar/profissão etc]

[-dora] = [nominalizador] [feminino] [agente]

[-nte] = [nominalizador] [sem gênero] [agente]

[-ista] = [nominalizador] [sem gênero] [agente]

O filósofo americano Fodor (*op.cit.*) radicaliza ao conceber o significado das palavras, no nosso caso aqui, o significado das entradas-sufixos, como átomos linguísticos, no sentido de que seus significados não se dividem, e rivaliza com as concepções contextuais e gradativas do significado.

Mas vamos continuar investindo na inclusão de potencialidades contextuais que contaminam as entradas-sufixos em estudo. Vamos assumir que o sentido é um *continuum*. Quando os contextos entram em jogo, algumas diferenças são maiores e abertamente mais claras, de que outras:

(127) Serafina considera o rapaz um *peniqueiro*.

(128) Serafina costuma ter ojeriza a *terreiro* de candomblé.

(129) Serafina vê um *abacateiro* no quintal do vizinho.

(130) Serafina contratou um *pipoqueiro* para a festa do netinho.

(131) Serafina gosta dos versos ‘que *braseiro*, que fornalha’.

É importante ser capaz de decidir se interpretações de uma entrada-sufixo para contextos diferentes representam uma ou mais de uma unidade semântica. Isso não é puramente uma questão teórica: por exemplo, um lexicógrafo terá que decidir quantas definições devem ser dadas à entrada-sufixo [-eiro], já que uma RFP que contenha uma entrada-sufixo como [-eiro] “é cega” em termos semânticos, ao se fixar na acoplagem estrutural tipo [nome + eiro].

Em princípio, podemos dizer que cada uma das entradas-sufixos [-eiro] nas expressões acima é uma unidade básica, nesses termos estaríamos concordando com a concepção de átomo linguístico *a la* Fodor. Poderíamos também afirmar que, adaptando a expressão texto-colônia, uma entrada como [-eiro] corresponde a uma entrada-colônia, que abrigaria variados traços semânticos, dada a identidade fonológica da entrada em questão em contraste com o seu feixe de traços semânticos.

Diante de um impasse como esse, impasse, diga-se de passagem, para um linguista, mas problema nenhum para os falantes, podemos nos posicionar dizendo que a unidade básica de significado da uma entrada-sufixo multifacetada semanticamente como [-eiro] é o *sentido*, e podemos dizer que uma entrada possui X sentidos se e somente se X contiver sentidos ambíguos. Precisamos agora ser mais explícitos no que queremos dizer quando afirmamos que uma entrada-sufixo é ambígua.

### 7.1.3 Ambiguidade das entradas-sufixos

Consideremos as entradas-sufixos [-eiro] nas sentenças:

(132) Serafina viu um *baleiro* no ônibus que vai para Placafor.

(133) Serafina considera o orientando um verdadeiro *peniqueiro*.

Poderíamos alçar o traço semântico avaliativo para *peniqueiro* e fazer o mesmo para *baleiro*? Ou seja, estaríamos avaliando a condição de classe social inferior das personagens em questão, ou simplesmente estaríamos descrevendo uma sena cotidiana de rua, sem valor

apriorístico pejorativo? No caso de *peniqueiro* não há evidências de traços semânticos em competição, dado que não existe a profissão de *peniqueiro*, mas no caso de *baleiro* há traços em competição, na medida em que existe a profissão, e contextos possíveis poderiam acionar a leitura avaliativa ou a leitura descritiva do personagem em questão.

Na ausência de um contexto pré-concebido, portanto, as duas leituras de *baleiro* estão em competição uma com a outra, ou, por vezes, atuam ao mesmo tempo: Serafina pode ter visto um *baleiro* e desdenhar, ao mesmo tempo, da profissão do rapaz.

Retomando uma metáfora de Cruse (*op.cit.*, 40): é:

(...) como as duas construções visuais de um cubo Necker<sup>24</sup>, somente uma pode ser o foco de nossa atenção num dado momento. Num contexto particular, um falante pretenderá dizer somente um dos sentidos e esperará que o ouvinte faça a mesma seleção. Não há um significado geral (...) que assuma as duas alternativas, ou opções de persistência não compreendidas, ou tomar ambos os significados de forma ampla: essas opções não estariam disponíveis (deixando de lado um deliberado jogo de palavras).

Cruse estabelece três testes para averiguar as condições de ambiguidade de entradas-palavras: *o teste de identidade*, *o teste das condições de verdade independentes* e *o teste zeugma*. Ele sugere que palavras ambíguas típicas passam por esses testes tradicionais de ambiguidade. Vamos aplicá-los a entradas-sufixos para observar a que conclusões podemos chegar.

TESTES DE AMBIGUIDADE		
TESTES DE IDENTIDADE	TESTE DAS CONDIÇÕES DE VERDADE INDEPENDENTES	TESTE ZEUGMA

Quadro 11

<sup>24</sup> Cubo desenhado pelo cientista suíço Luis Albert Necker (1786-1861), e publicado num trabalho de 1832, cujas linhas dão a ilusão de profundidade.



a) O teste de identidade

Em *Serafina viu um pipoqueiro*; e *Serafina chamou o bombeiro*, a entrada-sufixo [-eiro] pode ser interpretada da mesma maneira em ambas as ocorrências: trata-se da profissão de alguém. Isso mostra que [-eiro] compartilha um mesmo traço semântico para ambas as ocorrências, portanto, não estamos diante de ambiguidade aqui. Em comparação a: *Serafina desdenhou o baleiro* há uma pressão para interpretar a entrada-sufixo [-eiro] levando em conta seu traço profissão como também seu traço avaliativo. Já em *Serafina compadeceu-se do baleiro*, percebemos que o traço profissão foi alçado e o traço avaliativo foi apagado.

b) Teste das condições de verdade independentes

Vamos imaginar alguns contextos em que duas alternativas estão em jogo, dado o escopo sintagmático da entrada-sufixo em destaque abaixo:

(134) Serafina planta *bananeira* direitinho.

(135) Serafina *planta bananeira* na escola.

Isso mostra que a leitura com o escopo como em (134) e aquela com o escopo como em (135) – o objeto sendo tragado pelo verbo – têm sentidos distintos. Há, portanto, possibilidades simultâneas de que só um contexto extra-linguístico é capaz de desfazer. Ou seja, um acionamento do valor de verdade da sentença.

c) O teste zeugma

Perceba que os contextos em (134) e (135), para [plantar bananeira], acima, estão bem delimitados: em (134) [-eira] se refere a um pé de fruta, mas em (135), já que o verbo traga o objeto, acionamos o contexto, no nosso caso, o contexto sentencial como escrito acima, para desfazer a ambiguidade. Mas, simplesmente, se neutralizássemos [-eiro] em algo como:

(136) Serafina planta bananeira.

percebemos que os dois contextos foram ativados, portanto houve uma pane de sentidos. O efeito de pane para a ambiguidade acima é configurado por dois traços para a entrada-palavra serem alçados ao mesmo tempo. Essa não passou pelo teste.

Os testes explicitam que estão em jogo relações de sentido entre as entradas-sufixos e as entradas-bases às quais os sufixos se acoplam. Relações de sentido são relações de significados entre as entradas e seus elementos constituintes.

Cada entrada-sufixo mantém uma relação semântica de algum tipo com uma outra entrada, mas nem todas as relações possuem algum interesse intrínseco. Para ser de interesse,

uma relação precisa ser capaz de sustentar generalizações necessárias, que podemos classificar em paradigmáticas e as sintagmáticas. As relações sintagmáticas envolvendo o conceito de entrada lexical foram problematizadas no capítulo “A noção de entrada lexical e na estrutura argumental”, já que aqui estamos investindo em um recorte em semântica morfolexical, observando o comportamento semântico das entradas-sufixos.

As relações de sentido paradigmáticas são relações entre significados de entradas, que, nas palavras de Cruse (*op.cit.*, 68), *podem ocupar o mesmo rastro lexical*, e servem para unir a amplitude dos significados lexicais disponíveis num ponto particular da entrada vista como um todo.

Tome, por exemplo, a entrada [-eiro] em *Serafina chamou o pipoqueiro* em comparação com a ausência de afixo na sentença *Serafina viu a banan-*. Há um conjunto estruturado de escolhas de afixos para preencher o espaço vazio na sentença acima. Uma das escolhas pode ser por afixo do tipo derivacional como *Serafina viu a bananeira*, ou afixo de um outro tipo como *Serafina viu a banana*.

Podemos observar que as entradas-sufixos preveem uma articulação com uma certa configuração da sentença. Pois seria estranha a ocorrência em (137):

(137)?Serafina comeu a bananeira.

Mas nenhuma estranheza em relação a:

(138)Serafina comeu a banana.

Nesses termos, contestamos a assertiva corrente em variados trabalhos em morfologia (ROCHA, *op.cit.*; KATAMBA, 1992; SCALISE, *op.cit.*) de que a derivação não influencia no todo da sentença, papel da flexão. Veja que não existe uma influência estrutural da entrada [-eiro] em relação à sentença, mas existe uma influência semântica, dada a estranheza em (137) acima.

Vamos continuar investindo nas relações de sentido paradigmáticas aplicadas às entradas-sufixos, e considerar a divisão feita por Cruse (*op.cit.*) em dois tipos, de um lado, relações de *identidade e inclusão*, e, por outro, relações de *oposição e exclusão*.

RELAÇÕES PARADIGMÁTICAS	
IDENTIDADE E INCLUSÃO	OPOSIÇÃO E EXCLUSÃO

Quadro 12

## Relações de inclusão e identidade I: hiponímia

Começamos pelas relações de inclusão. Há um tipo básico dessas relações: a inclusão é de uma classe em outra, como no caso de [-eiro] e árvore, em que [-eiro] constitui uma subclasse incluída na classe mais ampla das árvores; assim como [-eiro] e profissão e [-eiro] e gentílico: [abacateiro], [bombeiro] e [brasileiro].

A relação inclusão-classe, chamada *hiponímia*, é exemplificada pelos dados que seguem: *coqueiro:árvore; faqueiro:objeto; jardineiro:profissão; brasileiro:gentílico*, e assim por diante; dos dois itens que mantêm relações entre si, o mais específico é chamado de hipônimo (é o caso de *coqueiro, faqueiro, brasileiro e jardineiro*, nos exemplos acima); e o mais geral é chamado de hiperônimo (ou superordenante, termo mais comum em língua inglesa), é o caso de árvore, objeto, profissão e gentílico, por exemplo. Note que, embora *coqueiro* seja hipônimo de *árvore*, esse termo é, digamos, um hiperônimo *coqueiro-anão*.

Um hipônimo pode ser pensado como uma relação do tipo “\_\_\_ é \_\_\_”, que pode preservar a verdade de declarações gerais tais como *Um coqueiro é uma árvore*, ou *Jardineiro é profissão*. Para um item lexical X ser um hipônimo de um item Y, a verdade de X ser Y advém, logicamente, dos significados de X e Y. A expectativa de que, se algo é X seja então bem provável que seja Y também, não é suficiente. Por exemplo, se alguém fala sobre uma barraqueira, muitas pessoas assumirão que a barraqueira em questão é alguém que se comporta de forma não aceitável. Entretanto, isso não nos autoriza a dizer que barraqueira é um hipônimo de profissão, porque há traços semânticos que podem ser acionados não relacionados com a profissão de alguém.

## Relações de identidade e inclusão II: metonímia

Segundo Cruse (*op.cit.*, 69):

A relação parte-todo, em seu aspecto lexical, é chamada de *metonímia* (algumas vezes de *partonímia*); por exemplo, *dedo* é uma metonímia de *mão*, e *mão* é um *holônimo imediato* de *dedo*. A noção de metonímia, como a de hiponímia, é muito mais relacional que absoluta: *mão*, por exemplo, é um holônimo de *dedo*, mas é, ao mesmo tempo, uma metonímia de *braço*, que por sua vez é uma metonímia de *corpo*. A série de relações pára em *corpo*, que pode ser chamada de *holônimo global*.

Outros exemplos de metonímia em termos de entradas-sufixos são os que se seguem: *bananeira: banana, pipoqueiro:pipoca, faqueiro:garfo, peniqueiro:penico*. Pares metonímicos prototípicos (em que X é metônimo de Y) são normais em composições como X

é uma parte de Y; um Y contém um X; as partes de Y são A, B, C... e assim segue. A metonímia pode ser claramente distinta da hiponímia, embora ambas envolvam espécies de inclusão. Uma maneira fácil de perceber a diferença é notar que uma banana não é um tipo de bananeira (metonímia), nem taxeiro é uma parte táxi (hiponímia).

#### Relações de identidade e inclusão

Vamos fazer algumas considerações sobre a noção de sinonímia e logo depois vamos observar como esse fenômeno de inclusão pode ser visto ao nível das entradas-sufixos.

Antes de um linguista sistematizar e dizer de uma maneira, digamos, formal, com um jargão próprio, usando uma linguagem acadêmica monitorada e repetida por seus pares, as estratégias de significação utilizadas pelos falantes para produzir e entender os significados com suas atitudes de falantes – os falantes já significam ou produzem significados por natureza, desde quando são bebês e disparam o processo de aquisição da linguagem (vimos isso em discussão anterior).

Os falantes, por exemplo, numa mesma situação de comunicação, podem dizer a mesma coisa utilizando palavras e frases diferentes. Os linguistas atribuem a uma estratégia como essa o nome de *sinonímia*. No dia a dia das nossas falas, nós nos damos conta de que podemos dizer a mesma coisa de maneiras diferentes, isso porque não estamos lá muito interessados, numa situação informal de comunicação (numa mesa de bar, por exemplo), com o que falamos, simplesmente, falamos o que temos a dizer.

Nesses termos, frases como as que seguem abaixo, em (a) e em (b), podem ser utilizadas uma pela outra sem que o seu conteúdo semântico seja alvo de um escrutínio científico para saber se, de fato, elas contêm conteúdo semântico equivalente.

a) Todo **careca** sonha com um remédio contra a calvície que dê resultados; e

b) Todo **calvo** sonha com um remédio contra a calvície que dê resultados.

Podemos perceber, à primeira vista, que essas duas frases podem ser utilizadas numa mesma situação sem prejuízo ou diferenças substanciais entre elas. O conteúdo semântico comum entre [**careca**] e [**calvo**] pode ser compartilhado num dicionário como o Aurélio. Se um leitor verificar esses verbetes num dicionário, ele possivelmente vai encontrar a palavra [**careca**] no verbete em que se encontra [**calvo**], e vice-versa. Há uma relação de sinonímia entre elas, ou seja, uma relação de proximidade semântica.

Segundo alguns estudos consensuais sobre a sinonímia, para que duas palavras ou expressões sejam sinônimas, não basta que tenham a mesma referência no mundo, ou seja, referir-se a um mesmo objeto no mundo não é suficiente para que palavras e expressões sejam

sinônimas. Para que sejam sinônimas, além de se referir a uma mesma entidade no mundo, elas terão que ter o mesmo sentido.

O que é ter o mesmo sentido? Saber o sentido de uma palavra ou frase é ser capaz de dizer se, em determinadas circunstâncias, ela é verdadeira ou falsa. Duas frases têm o mesmo sentido quando se referem ao mesmo conjunto de fatos no mundo, e tem de ser ambas verdadeiras ou falsas. Como não houve alteração entre a verdade e a falsidade nos conteúdos das sentenças em (a) e (b), [*calvo*] e [*careca*] são sinônimos para essas duas sentenças. Entretanto, podemos simular uma determinada situação em que isso não ocorre:

c) Benício não se chateia quando é chamado de *calvo*, mas não suporta ser chamado de *careca*; e

d) Benício não se irrita quando é chamado de *careca*, mas não gosta de ser chamado de *calvo*.

Dadas simulações como essas em (c) e em (d), os semanticistas se dividem basicamente em três concepções sobre o fenômeno da sinonímia: há aqueles que dizem que sinônimos não existem numa língua; há aqueles que defendem a não existência de sinonímia perfeita; e há ainda outros que preferem não abandonar o conceito de sinonímia, dado que ele é produtivo na língua, enquanto flagram os falantes dizendo praticamente a mesma coisa com palavras e frases diferentes numa mesma situação comunicativa.

Não podemos considerar nas sentenças (c) e (d) que [*calvo*] e [*careca*] sejam sinônimos, pois não gozam de um mesmo sentido. A conclusão parcial a que podemos chegar é que não é possível considerar sinonímia fora de um contexto, assim como não é possível considerar a sinonímia deixando de lado a noção de composicionalidade das palavras e frases.

Segundo os filósofos americanos Katz e Fodor (*op.cit.*), composicionalidade é uma estratégia de uso da língua do ponto de vista da mente humana que basicamente prediz que a composição de uma sentença é uma combinação entre suas palavras numa frase, uma espécie de fricção, de um roçar de alguma maneira entre elas, que, ao se friccionar umas com as outras produzem sentidos diferentes, pois “os falantes não estão diante de elementos novos, mas diante de uma nova combinação de elementos conhecidos”.

Consideremos as frases que seguem:

e) Andreza chamou o *menino* para lavar as mãos.

f) Andreza chamou o *garoto* para lavar as mãos.

g) A velhinha telefonou para um *garoto* de programa.

h) A velhinha telefonou para um *menino* de programa.

Podemos considerar a sinonímia em seu valor conceitual – nesses termos, [*menino*] e [*garoto*], e [*calvo*] e [*careca*] se equivalem. Mas, se levarmos em conta estilo, contexto, associações sociais e dialetais (sotaques regionais), ou registros (falas individuais: as de mulher são diferentes das de homem, por exemplo; assim como as dos infantes são diferentes das dos adultos) só podemos considerá-la em seu aspecto proposicional (quando estão inseridas numa frase), e, inseridas numa frase, o seu valor de verdade ou falsidade é alterado.

*Gordo* e *obeso*, por exemplo, somente podem ser intercambiáveis em determinados contextos. *Gordo* em contextos mais amplos, e *obeso* em contextos mais estritos, por um endocrinologista, por exemplo. Intercambiar [*garoto*] e [*menino*] em (g) e (h) é praticamente impossível ou inviável, mas já em (e) e (f) eles podem ser trocados um pelo outro.

Nesses termos, podemos falar em sinonímia gradual, ou seja, mesmo consideradas sinônimas, as palavras sofrem algum tipo de especialização no uso. Quando dissemos logo acima que os falantes significam por natureza, queremos dizer que estratégias de significação são coisas tácitas – dadas como naturais – para os falantes. Para os linguistas, no entanto, isso parece ser um problema quase insuperável, como demonstramos nesses poucos exemplos envolvendo o fenômeno da sinonímia.

Mas a sinonímia não é restritiva o suficiente, para se aplicar com certo grau de certeza a todos os casos, como [-dor] e [-eiro], como em jogador e ciclista, ambas as entradas [-dor] e [-eiro] podem se referir a atividades, alguém fazendo algo, a noção de agente. Seria útil, entretanto, examinar mais cuidadosamente a noção de *mesmo significado e significado similar*.

Pares ou grupos de sinônimos podem ser caracterizados de acordo com quão próximos os significados das entradas são. Três graus de proximidade podem ser reconhecidos: sinonímia absoluta, sinonímia proposicional e sinonímia de proximidade.

A maior semelhança possível entre dois sentidos é a identidade, em outras palavras, de sinonímia absoluta. Vamos recorrer às idéias de Haas (citado por Cruse, *op.cit.*) para problematizar os graus de proximidade acima.

Em todos os contextos (gramaticalmente bem formados) que preservem uma igual normatibilidade, estamos diante de possível sinonímia. Isso se baseia na assertiva de que qualquer diferença de significado revelará por si mesma uma diferença de possibilidades de co-ocorrência, daí a descoberta de um contexto em que um dos sinônimos supostos é mais normal que o outro e regule o par como sinônimos absolutos.

Esse é um critério extremamente restrito, e um teste rigoroso dos pares candidatos que levem rapidamente à convicção de que sinônimos absolutos sejam difíceis de ser encontrados em termos de entradas-sufixos. Como os exemplos abaixo demonstram:

- (a)ouvidor/ouvinte
- (b)batedora/batedeira
- (c)inventador/invencioneiro
- (d)falante/falador

Esses pares não são sinônimos absolutos por diferenças de normatibilidade entre (a); (b); (c) e (d):

- a) Ele é o ouvidor.
- b)Ele é ouvinte.
- c)Ele é um falador.
- d)Ele é um falante.
- e)Ele é um inventor.
- f)Ele é um invencioneiro.

O que Haas compreende como normatibilidade é uma leitura particular de um item em relação ao seu ambiente de ocorrência. Veja que tanto [-eiro] como [-dor] acionam o traço agente, mas nas ocorrências acima, esse traço não pode ser aplicado a todos os pares em questão em que ele ocorre.

A sinonímia absoluta está mais próxima presumivelmente daquilo que as pessoas têm em mente que as faz acreditar em sinônimos, e geralmente é pouco provável que ela ocorra em línguas naturais. Procuramos um par que explicita sinonímia absoluta e não encontramos. Nesses termos, os contextos ajudariam a marcar a sinonímia como absoluta.

Há talvez um ou outro caso, por assim dizer, de sinonímia absoluta cuja identidade de significado pode ocorrer entre formas que pertençam a diferentes variedades, especialmente diferentes dialetos de uma língua. Um exemplo óbvio pode ser *fall e autumn* (outono), empregados no inglês americano e no inglês britânico, respectivamente. Não há diferença em princípio para equivalentes transnacionais em outras línguas. Note, entretanto, que esse dado não poderia ilustrar uma sinonímia absoluta pelos testes de Haas, já que *fall* seria menos normal que *autumn* num dado contexto sentencial que pode ser, por outro lado, lexicalmente marcado como inglês britânico. Assumir que *fall e autumn* são idênticos em significado pressupõe uma noção não-Haasiana daquilo que significado quer dizer.

## Relações de oposição e exclusão

Segundo Cruse (*op.cit.*, 74):

Oposição e sinonímia são provavelmente as únicas relações de sentido familiares às pessoas comuns. A maioria das línguas possui palavras habituais para as relações de exclusão; é uma relação cognitivamente muito básica e quase sempre todas as crianças já em tenra infância podem compreender essa noção. Oposição é um tipo especial de incompatibilidade: pares opostos são inerentemente binários, ou seja, ambos aparecem natural e logicamente em pares.

Não vamos explorar tais relações aqui, somente queremos registrar que alguns sufixos (os quais elencamos acima) não mantêm relações simétricas entre si.

Resumo: Discutimos o estabelecimento do conceito de entrada lexical ao nível morfolexical, e fixamos o conceito de entrada-sufixo, tendo a derivação como os pressupostos teóricos que problematizaram essas questões.



## **8 A CHECAGEM DAS ENTRADAS LEXICAIS**

FOCO: Este capítulo vai tentar estabelecer bases na investigação de como a checagem das entradas lexicais é feita do ponto de vista cognitivo, ou seja, da mente humana de quem fala para fora da mente humana de quem fala.

Ao construir sentenças, os falantes não usam todos os itens que estão em suas cabeças ao mesmo tempo nem tampouco lançam mão de todas as estruturas que dominam em sua língua para poder por o que está em suas mentes em uso: os falantes fazem escolhas na organização das frases.

No uso, do ponto de vista discursivo, talvez seja mais viável determinar quais escolhas, que tipo de escolhas, quais estratégias, suas prioridades, os falantes são levados a fazer. Por exemplo, podemos dizer que os falantes, de uma forma ou de outra, levam em consideração ambientes contextuais formais e não formais, e depois falam o que têm a falar. Podemos dizer que os falantes imaginam quem são seus interlocutores, e, a partir disso, dizem o que têm a dizer. As escolhas, portanto, nesses termos, estão condicionadas a critérios contextuais.

Diante das hierarquias sociais, do grau de parentesco, do grau de proximidades e intimidades, os interlocutores, de uma forma ou de outra, escolhem algumas frases e não outras, algumas palavras e não outras, um modo de dizer e não um outro. Os falantes jogam o jogo dos rituais sociais de fala.

O problema grave, contudo, pois antes de as palavras saírem da boca de alguém, elas estão no cérebro de alguém, é determinar se há escolhas na mente antes do uso. Se houver essas escolhas, portanto, como elas são feitas? Para introduzir um termo técnico, vamos

assumir que essas escolhas – do ponto de vista da mente do falante – são estratégias cognitivas, e deixar de lado duas coisas:

- a) a noção de escolha e introduzir a noção de checagem; e
- b) a noção de palavra e de afixo e, em seu lugar, lidar com a noção de entrada lexical.

Vamos reformular as bases teóricas deste capítulo, assumindo a questão – em consonância com o proposto por Aronoff (2001, p.239), repetido aqui: “...o falante de alguma maneira checa o léxico, para ver se uma palavra está lá...” –: quais estratégias cognitivas os falantes empreendem na checagem do léxico? Ou seja, como é que, “de alguma maneira”, a checagem chega a termo.

A checagem é tácita, subjacente, e impossível de não ser feita, porque os falantes, em meio ao léxico de que dispõem em suas mentes, escolhem certas palavras e não outras, escolhem alguns padrões e não outros, assumem algumas estruturas e não outras.

Este capítulo vai tentar estabelecer bases na investigação de como a checagem é feita do ponto de vista cognitivo. Pois, “o mundo das palavras é tão maravilhoso quanto o mundo da sintaxe ou até mais”. (PINKER, 1994, p.152)

A mente é um universo e os astros são as entradas lexicais. Na viagem de uma entrada lexical pelo universo da mente, saindo do dicionário mental até a estrutura de uma sentença, os falantes asseguram que as entradas cheguem “inteiras” ao seu destino final – uma sentença. E elas chegam inteiras pois eles acertam todas as vezes em que constroem sentenças em sua língua natural: do léxico para a estrutura, a viagem é sempre bem sucedida. As questões que serão tratadas aqui basicamente cobrirão assertivas bem conhecidas como: os afixos têm posição diferente da base na jogo com as entradas lexicais; uma entrada lexical atua em um domínio duplo, por exemplo: há entradas que atuam no léxico (afixos e bases), há entradas que atuam na estrutura sentencial (as palavras). Afixos, bases e palavras são designados aqui como entradas, pois todos eles sofrem leitura semântica. Ou seja, se uma palavra pode significar algo, um prefixo ou um sufixo pode significar algo também. Assim, se uma palavra como “maçã” pode significar ou pode ser uma fruta; um prefixo como “des-” pode significar negação.

Vamos considerar que, para organizar o seu dicionário mental e a checagem das entradas na viagem do léxico para a estrutura, o falante lance mão da morfologia e do léxico. Uma representação desses dois aspectos – morfologia e léxico – na checagem proposta envolve um conjunto de operações que, por não termos palavra melhor, vamos designar como gerativa: morfologia e léxico geram as entradas para a estrutura.

Vamos concordar, portanto, que esses dois níveis atuam em interface: enquanto um nível – a morfologia – atua por intermédio de padrões; o léxico, por sua vez, atua de maneira solitária e cumulativa: as entradas são acumuladas na mente do falante, e armazenadas na memória.

Vamos admitir também que os falantes necessitem de informações semânticas para essas entradas na hora da checagem, e que as propriedades semânticas das entradas são propriedades descritas em termos de traços semânticos.

Para não ser acusado de vaguides, vamos definir um traço semântico de uma entrada em termos bem informais: um traço semântico é uma nesguinha de significado contida numa entrada. Por exemplo, para a entrada [banco] digamos que podemos pontuar várias nesguinhas de significado: instituição financeira, móvel de sentar, obstáculo para os navegantes num barco (banco de areia) etc. Cada uma dessas nesguinhas, portanto, um traço semântico. Como os falantes não usam esses traços todos ao mesmo tempo, para usar essa entrada, eles checam o traço que pretendem, alçam-no num processo de projeção semântica (ver capítulo 4 para lembrar a noção de projeção semântica), e elegem o traço necessário que caberá na sentença que pretendem pronunciar.

Podemos afirmar, sem nenhum receio de estarmos equivocados, que uma entrada lexical possui traços semânticos, e gramaticais, que também são checados para a viagem das entradas da mente para fora da mente humana. Traços semânticos lidam especificamente com o que uma entrada significa, e traços gramaticais distribuem as palavras em classes gramaticais ou categorias gramaticais.

Por exemplo, [banco] terá os traços semânticos mais ou menos como os descritos acima, e traços gramaticais mais ou menos como os descritos no quadro abaixo :

VERBO	Eu banco aquela mulher, e ela nem me retribui.
NOME	O banco de areia atrapalhou o retorno do navio.
NÚMERO	A entrada lexical banco está no singular.
GÊNERO	Banco é uma entrada masculina em português.
SUJEITO	Na sentença ‘o banco é grande’, ‘banco’ é sujeito.
OBJETO	Na sentença ‘ela viu o banco’, ‘banco’ é objeto direto.

Quadro 13

Vamos introduzir uma outra noção técnica que acreditamos envolvida na checagem: *o princípio da interpretação plena*.

Basicamente, queremos dizer que a checagem é feita envolvendo ao mesmo tempo traços semânticos e traços gramaticais. Isso significa que, no deslocamento da entrada do léxico para a sentença, a interpretação plena especifica que a representação de uma dada entrada na estrutura contém todos os elementos envolvidos do ponto de vista gramatical e semântico necessários para que a entrada dê certo na estrutura.

A interpretação plena contribui para isso graças ao envolvimento da morfologia e do léxico. Por exemplo, a morfologia vai acionar o traço pessoa para a entrada [banco] na sentença:

- a) “Banco aquela mulher e ela não me retribui.”

Ou o traço número nas sentenças:

- b) “Banco: instituição financeira para rico.”  
c) “Bancos: instituições financeiras para rico.”

Em (a) poderemos ter o traço número também, mas é acionado um traço gramatical de outra espécie.

O princípio da interpretação plena aciona, portanto, morfologia e léxico, e opera em conjunto com esses dois níveis para disponibilizar uma entrada para a sentença: não há vida para uma entrada lexical na estrutura se não houver a interferência desses dois níveis ou sistemas da língua.

Quando uma representação de uma dada entrada satisfaz de forma plena o sistema morfológico, está-se cobrindo o nível gramatical; se uma representação satisfaz o sistema lexical, está-se cobrindo o nível semântico. Ambos os sistemas devem ser satisfeitos sob o risco de haver um *crash* ou uma *pane* na formulação de uma sentença. *Crash* significa o acionamento inadequado desses dois sistemas, como a anomalia sentencial abaixo pode deixar isso explícito:

- a) \*“Eu banco de areia aquela mulher e ela não me retribui.”

Nessa pseudo-sentença imaginária acima, a morfologia acionou traços gramaticais em total desacordo com os traços lexicais exigidos para a representação ideal, pois houve uma espécie de *crash* entre morfologia e léxico: a checagem dos traços trocou as bolas. Num lugar em que o traço semântico lexical acionou nome, a morfologia acionou verbo para o item [banco], por isso a sentença não é somente estranha, como também anômala, e impossível de ser salva semanticamente, como fizemos ao salvar semanticamente no capítulo VIII o sintagma [cachorro falante]. Veja que [cachorro falante] pode ser salvo semanticamente, já

que não houve *crash* entre morfologia e léxico. Nesses termos, a sentença (a) acima é totalmente nada gramatical, já que a checagem não foi feita como propomos nessa linha de argumentação.

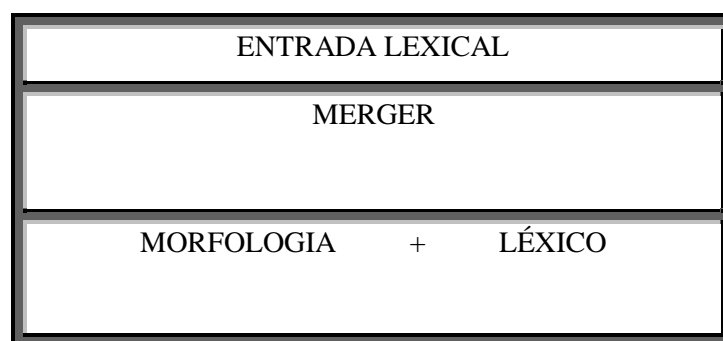
A interpretação consequente do princípio de interpretação plena que satisfaça a esses dois sistemas deve conter o acionamento necessário, e ao mesmo tempo, desses dois sistemas, sob o risco de as entradas serem algo inexistente na língua. Isso sugere que no caminho em direção a uma sentença bem formada, esses dois níveis ajam de forma a satisfazer um equilíbrio lexical e morfológico.

No caminho dos astros-entradas lexicais para a estrutura-sentença há seleção rigorosa de traços que façam parte desses dois sistemas.

Vamos tomar emprestado da sintaxe gerativa a noção de *merger*, uma espécie de fusão, incorporação, constituindo uma combinação de traços. Na fusão de traços necessariamente adequados, um conjunto de ataques, ou atributos gramaticais e lexicais de traços, que estão em lugares mais ou menos separados na mente humana, precisam de uma operação de *merger* para que se consiga juntar os traços certos para desembocar no uso a sentença sadia, sem nenhum tipo de doença gramatical.

Quando firmamos a hipótese de que esses dois conjuntos estão separados na mente humana, é porque, como pontuamos logo no primeiro capítulo, léxico e morfologia podem se impor como dois sistemas em separado, apesar de agirem como uma equipe ao lidar com as entradas lexicais.

Vamos representar esse modelo como no esquema abaixo:



Quadro 14

O quadro acima demonstra que a atuação dos dois sistemas conflui para que uma entrada lexical esteja disponível para entrar numa sentença (no caso de uma palavra), ou esteja pronta para a formação de uma palavra (no caso de afixos).

Vimos que, da combinação e/ou fusão de traços, estabelecem-se determinadas escolhas (checagem) e não outras. Uma questão que surge, imprescindível para a continuidade dos processos de checagem, é saber como alguns traços são alçados e outros ficam sem participar da sentença. Ou seja, como assegurar que os traços certos são aqueles associados à sentença correta. Aqui reside a questão central de nossa investigação sobre checagem.

Para mapear a checagem com acuidade, vamos observar como ela funciona para três tipos de entradas lexicais numa língua: as entradas lexicais flexionais; as entradas derivacionais; e as entradas quando na estrutura da língua, ou seja, uma palavra. Para isso, vamos lançar mão de pressupostos da hipótese lexicalista para os três tipos de entradas em foco.

## 8.1 A LÓGICA DA FLEXÃO

Entradas flexionais numa língua desempenham papel fundamental na estrutura sintática, como também morfológica, para outras entradas e para uma sentença, e pertencem a uma categoria particular. Possuem traços gramaticais e semânticos e se combinam com outras entradas para determinar um sentido também particular. Atente-se aos exemplos abaixo:

- a) “Mesa-[s] geralmente **são** feitas de madeira.”
- b) **Nós** come-[mos] frutas saudáveis.”

Os itens entre colchetes têm algo a ver com os itens em negrito. Há uma combinação entre eles que podemos chamar de concordância, em outros termos, há uma relação sintática necessária entre eles sob pena de, na ausência dessas entradas flexionais, a sentença parecer estranha, não gramatical de alguma maneira. A ausência ou presença desse tipo de entrada flexional determina a boa formação de uma sentença.

A identidade de sentenças ou de uma entrada flexional varia de forma sistemática de acordo com o contexto particular onde ela esteja ou não esteja inserida. Em português, uma entrada lexical como [cantar] pode ser estendida para entradas diferentes, conforme a acomodação de itens flexionais: cant-o / cant-a-mos / cant-ei / cant-ou / cant-a-ria. O tipo de entrada flexional acoplada à base do item é fundamental para determinar tempo, pessoa e número. A checagem de se o falante quer dizer algo no passado ou no futuro depende de escolhas que ele fizer em termos de entradas flexionais para o verbo em português. Não estamos dizendo nenhuma novidade com isso. Pois a divisão das entradas em flexionais ou de outro tipo é um dos achados mais antigos em linguística, que remonta aos gregos.

Na checagem de como as entradas como palavras se realizam na mente dos falantes – os falantes retiram de um lugar da mente uma entrada flexional e acopla essa entrada a uma base para completar, dar novo sentido, estabelecer categorias como tempo e pessoa –, a morfologia entra no jogo constituindo *paradigmas* flexionais.

A estrutura de paradigmas numa dada língua é determinada pelo inventário de propriedades morfossintáticas disponíveis para essa língua. Os falantes dominam isso de uma forma tão perfeita e automática que uma entrada como [-s] pode fazer parte de paradigmas diferentes, nominais ou verbais. Da fricção entre essa entrada e uma base, o paradigma distribui as palavras em nomes ou verbos. Como as entradas abaixo podem ilustrar isso:

- a) “Cadeira-[s] servem para sentar.”
- b) “Canta-[s] e teus males espantas.”

Conclusão provisória 6: distribuição em paradigmas, ou seja, a ação da morfologia, é acionada pelos falantes na checagem das entradas em nominais ou verbais.

Temos um problema aqui. Há uma espécie de fricção entre as entradas [cant-] e [-o], e não sabemos a priori de que paradigma gramatical a entrada [canto] faz parte. A priori, [canto] é nome ou verbo? Como as entradas lexicais atuam de forma dupla, tanto no léxico como na estrutura, será necessário que essa entrada [canto] seja alçada a uma estrutura para que ela seja lida como nome ou verbo. Em (a), abaixo, nome; em (b), abaixo, como verbo:

- a) “O canto lírico é sublime.”
- b) “Canto para Joana com a minha voz de soprano.”

Conclusão provisória 7: o deslocamento do léxico para a estrutura completa o percurso das entradas.

Para estarmos completamente certos de que a checagem chegue a bom termo, sempre e necessariamente, e levando os dados (a) e (b), acima, em consideração, devemos separar flexão (presente em [a] e [b]), acima, de derivação, geralmente acionada na criação de novos itens na língua, e que não se relaciona diretamente com a sintaxe da língua – papel da flexão. A derivação foi problematizada no capítulo IX, por isso, não nos damos o trabalho aqui de explicitar mais sobre ela.

Nesses termos, dois diferentes tipos de morfologia estão disponíveis para a checagem dos itens na mente. De um lado, há mecanismos morfológicos que devem ser usados para deduzir se uma palavra pertence a este ou àquele paradigma. Por exemplo, uma entrada como

(-o) acima distribui a entrada (cant-) em nome ou verbo: processos de flexão atuando na separação entre paradigma verbal ou nominal.

Por outro lado, mecanismos morfológicos podem ser usados para deduzir se estamos diante de uma nova entrada na língua. É o caso do prefixo (des-), que não aparece em (a) abaixo, e que aparece em (b) abaixo. Sua ausência ou presença é fundamental para que o falante tenha a certeza de que ele está diante de uma entrada “nova” na língua.

- a) “Obediência é algo necessário para as leis.”
- b) “Des-obediência é algo fundamental para melhorar uma lei.”

Mas, diante de dados como (desandar) e (desferir) como os falantes sabem que o (des-) de (desandar) não necessariamente tem a mesma coisa a ver com o (des-) que atua em (desobediência)? Ou o (des-) de (desferir) não é a mesma coisa que o (des-) de (desfazer)? Se os falantes jogassem empiricamente todas as vezes em que checam as entradas, ou seja, se a combinação de afixos como entradas fosse apenas matemática, ou em cadeia, haveria um *crash* na atuação dos processos de flexão e de derivação, e, por conseguinte, nos processos de saída dos itens da mente para o uso: os falantes atuam por traços semânticos abstratos e não necessariamente por meio dos itens mecanicamente. As peças não se encaixam, portanto, mecanicamente. Elas se encaixam, digamos, analogicamente, resultando resultados por vezes imprevisíveis.

Nesses termos, o processo de interpretação plena joga com esses dois tipos de morfologia para saber acionar o traço certo na sentença certa. Uma frase de Chomsky (1995, p.58), que se vira para dentro da mente pra saber como a língua funciona, é mais do que uma frase de efeito, é algo quase que místico em se tratando de língua: “As palavras não existem, o que existe são traços abstratos”. Parafraseando para nossos propósitos, as entradas não existem o que existe são leituras de traços gramaticais e semânticos para as entradas que materializamos foneticamente no uso.

Para desdobrar os dois tipos de morfologia que propomos e investigar como esses dois domínios atuam nos processos de checagem, vamos estabelecer alguns critérios empíricos na velha e sempre retomada, mesmo que polêmica, separação entre flexão e derivação. Como o nosso propósito aqui é argumentativo, não vamos levantar uma série exaustiva de dados para isso, pois queremos ver como a mente funciona para esses dois tipos de morfologia.

## 8.2 CRITÉRIOS DE DISTINÇÃO ENTRE DUAS MORFOLOGIAS: A MORFOLOGIA DA FORMAÇÃO DE PALAVRAS E A MORFOLOGIA DOS PARADIGMAS FLEXIONAIS



Se (desfazer) não significa o mesmo que (fazer) e (comi) não significa o mesmo que (comer), por que, então, esse par de entradas é classificado em critérios diferentes, ora como morfologia de formação de palavras (fazer/desfazer), ora em paradigmas flexionais (comer/comi)? Já podemos ver que a noção “traços semânticos” não vai nos ajudar muito nisso. Deveremos partir para critérios, digamos, mais empíricos, na distinção desses dois tipos de morfologia.

Revirando uma literatura quase que já saturada na distinção entre esses dois tipos de morfologia, vamos retomar critérios estabelecidos por Câmara Jr. (1970), Rocha (1998) e, mais recentemente, Stump (2003), para fixar essa distinção, e associá-la a mecanismos mentais de checagem.

Comumente, cinco critérios são geralmente associados à distinção entre morfologia de paradigmas flexionais e morfologia de formação de palavras. Os critérios são independentes uns dos outros, mas todos eles jogam com traços semânticos e projeção de itens na estrutura da língua. Vamos manter a terminologia utilizada por Stump (*op.cit.*) por acreditarmos ser a mais criteriosa e abrangente, em comparação à utilizada por Câmara Jr. (*op.cit.*) e Rocha (*op.cit.*). Mas vamos adaptá-la aqui com mudanças para os nossos propósitos cognitivos, pois os critérios de Stump são empíricos, enquanto queremos nos distanciar o mais que pudermos de relações funcionais, empíricas e de uso.

- a) O critério da mudança de sentido lexical ou de parte do sentido lexical de uma entrada lexical.

Duas entradas lexicais que se relacionam por princípios de derivação devem se diferenciar de seu sentido lexical básico, ou não compartilhar sentido lexical básico em todo ou em parte. Duas entradas lexicais que fazem parte do sistema flexional compartilharão necessariamente um sentido básico. O que significa dizer que as diferenças de comportamento gramatical não serão capazes de afetar suas propriedades morfossintáticas, e que essas propriedades serão capazes de acumular um mesmo conjunto de células num paradigma.

Por esse critério, o padrão de formação de agentivos em português (acionado por afixos como [-eiro] e [-ente]) faz parte da morfologia de formação de palavras. Enquanto o padrão de número singular e plural [-o] e [0], presença ou ausência de morfema, deve fazer parte do paradigma flexional.

O diagnóstico estabelecido no primeiro critério acima depende da precisão com que podemos articular princípios para determinar se uma entrada compartilha sentidos lexicais em

termos paradigmáticos (portanto pertencentes a uma mesma célula lexical) ou não compartilha sentidos em comum (portanto, o lado derivacional da morfologia).

Esse critério sozinho não consegue dar conta do problema, apesar de sua lógica na distinção desses dois tipos de morfologia, por duas razões, propostas por Stump (*op.cit.*, p. 15):

- 1) “A mudança de sentido da entrada lexical nem sempre vem acompanhada da mudança de parte da entrada lexical”.

Como vimos isso no jogo que fizemos entre (desfazer/fazer) e (comer/comi). Ou seja, tanto na derivação quanto nos paradigmas flexionais, partes da mesma entrada sofrem recombinações, ou seja, (fazer) se combina com (desfazer), de algum modo, e (comer) se combina com (comi), de algum modo também. Assim, a simples combinação de elementos não necessariamente distingue os dois sistemas de morfologia. Nesses termos, o critério (a) é insuficiente.

- 2) “Pares homônimos e/ou sinônimos confundem o critério estabelecido acima”.

É o caso de (canto, nome) e (canto, verbo). Há uma necessária mudança de sentido para que a distinção se efetive. Nos dicionários que circulam por aí, como Houaiss e Aurélio, a distinção é marcada através de verbetes que distinguem canto/nome de canto/verbo. E veja que o que serve para a morfologia derivacional é o lado “nome” da entrada lexical (canto), e não o seu lado “verbo”.

O critério da ação da sintaxe sobre a entrada lexical.

O contexto sintático de uma entrada lexical faz com que ela se realize de uma forma particular em termos de paradigma flexional ou em termos de formação de palavra. (Canto) é um caso bem típico em língua portuguesa. Vamos adotar a categoria número como forma de separar os dois tipos de morfologia. Em termos de paradigmas flexionais, a categoria número se realiza como se segue:

- 3) “Cantamos no jardim.”

Se a mente aciona uma entrada lexical em termos derivacionais, a categoria número se realiza como se segue:

- 4) “Os cantos da sala estão limpos.”

E o que isso tem a ver com a checagem na mente do falante?

Vamos nos imaginar numa terra estrangeira, onde se fala uma língua estranha, da qual não temos a menor ideia de onde começa e termina uma palavra, não sabemos como se chama nem ao menos o nome de um líquido chamado água. Pois bem. Os primeiros contatos das crianças com a língua de seus pais simulam a entrada numa terra estrangeira, de língua

desconhecida. Mas, ao invés de entrarem em pânico, as crianças, fascinadas, se jogam no empreendimento de domínio dessa língua.

Elas já trazem um triunfo herdado da evolução: elas têm, digamos, um dispositivo em suas mentes para poder operar a empreitada de rapidamente e uniformemente, num curto período de tempo, saber que cada coisa tem um nome, e saber que podem construir frases com os nomes das coisas que têm um nome. Elas se jogam na tarefa de dominar aquele ambiente, dominar aquele ambiente desconhecido é dominar a língua que se fala naquele ambiente, através do domínio das entradas dispostas para aquela língua.

São muitas informações inicialmente, o mundo da língua é quase infinito. Existem milhares de coisas que devem ser aprendidas, milhares de nomes, toda frase que ela ouve é sempre nova, e a complexidade dos significados das palavras exige estratégias sofisticadas para o domínio desses significados. Como ela vai saber, por exemplo, que a palavra “chão” não designa somente o lugar onde ela pisa, mas algo como situação de inferioridade, o lugar onde alguém nasceu, o fundo de um quadro, piso, solo etc?

Vamos começar por assegurar que as palavras contenham traços semânticos, e vamos assumir que a palavra “chão” contenha os traços que delineamos: solo, situação de inferioridade etc. As crianças começam a tarefa de domínio dos significados das palavras traço a traço, ou seja, elas primeiro ficam sabendo que “chão” significa o lugar onde ela pisa, depois, suas experiências vão descobrir que não significa somente isso, e pode significar outras coisas também, como situação de inferioridade etc. Elas vão ficar sabendo que há determinados “chãos” onde ela pode pisar, e há outros dos quais ela deve se esquivar (o “chão” onde se encontra uma poça de lama, por exemplo).

Essa tarefa de domínio de traços contidos nas palavras empreendemos por nossa vida afora. Mas isso não responde ainda a questão posta logo acima. Pois, memorizar traços semânticos de palavras é algo secundário na apreensão dos significados e das entradas. Nossa memória é muito limitada, e não podemos confiar apenas nela para lembrarmos o significado das palavras com as quais entramos em contato. A apreensão dos significados não se dá através de uma lista imensa que jaz em nossas mentes, como se fosse uma lista telefônica. Nossa memória entraria em pane. Entra em cena algo que podemos assumir como o aparato mais sofisticado que temos para a apreensão de significados: o jogo com os padrões.

Como nós os adultos nem precisamos recorrer ao dicionário para estabelecermos o significado uma palavra nova? Suponhamos que a palavra “luleiro” apareça de repente na

língua portuguesa. E apareça na seguinte frase: “O manifestante luleiro aplaudiu as palavras do Presidente no comício em Brasília, os adversários vaiaram”. Veja que estou simulando o aparecimento de uma palavra nova para o mundo adulto, mas compreenda que para as crianças todas as palavras são “novas”, e elas têm a desvantagem de não poder contar com dicionários para socorrê-las na tarefa de apreensão de significados: elas terão que acionar o seu instinto, suas intuições e sua memória ao mesmo tempo para ir adquirindo e armazenando os significados em suas mentes.

Há uma estocagem de significados que se estabelece a partir do jogo com os padrões linguísticos e com o jogo das situações de fala onde elas ouvem essas palavras no seu cotidiano. As situações de fala são analisadas pelas crianças por intermédio de um viés de exclusão mútua. Imaginemos a seguinte situação: há um gato muito próximo de um cachorro. São dois animais distintos com nomes distintos. Descartamos a memória como algo secundário na apreensão dos significados. Como a criança será capaz de distinguir esses dois animais e entender que cada um deles tem um nome diferente um do outro?

A tarefa de exclusão mútua ajuda nessa distinção da seguinte forma: sons diferentes são assumidos para referentes (nesse caso, animais) diferentes. Mas a exclusão mútua ainda é um passo inicial no aprendizado dos significados. As crianças vão se deparar com animais grandes (um são bernardo) ou pequenos (um basset) e todos eles vão estar incluídos na família dos caninos, e, por conseguinte, receberão o nome de cachorro. Ela também pode se confundir em achar que o nome “Domane” (o nome do cachorro da família) designa o animal, inclusive, todo cachorro que ela encontra pela frente ela aponta e diz “Domane”. Como desfazer essa confusão de nomes, classes, famílias e traços semânticos?

As crianças contam com um outro viés no aprendizado dos significados: o amadurecimento do dispositivo linguístico herdado da evolução. À medida que o dispositivo amadurece, noções mais abstratas como inclusão de animais em famílias (o caso de cachorro, em basset, são bernardo, pudor etc) vão compondo as possibilidades de apreensão dos significados, e tornado-os mais abrangentes. Os significados vão se acumulando na mente das crianças não como uma lista infinita: mas com a organização das informações por meio de exclusões e padrões tendo o dispositivo da mente no comando. Para darmos continuidade ao trabalho de checagem, vamos observar duas possibilidades das entradas do ponto de vista semântico: a extensão e a intensão do significado das entradas.

CHECAGEM SEMÂNTICA	EXTENSÃO
	INTENSÃO

Quadro 15

As possibilidades e impossibilidades de colar o significado de uma palavra ao seu referente, ou seja, de saber que o objeto com encosto onde sentamos se chama “cadeira”, impõem uma distinção entre extensão e intensão do significado. Enquanto a extensão diz respeito ao conjunto de entidades que seleciona ou escolhe uma palavra para nomear um objeto, sua intensão corresponde ao seu sentido inerente, aos conceitos que ela evoca.

Nesses termos, a extensão da palavra “mulher” circunscreve um conjunto real de entidades (todas as mulheres), enquanto sua intensão envolveria noções como “fêmea” e “ser humano”; até mesmo coisas aparentemente contraditórias como macho: “Joana é uma mulher macho”. Vamos delinear no quadro abaixo o jogo das extensões e intensões que circunscrevem um área semântica:

FRASE	EXTENSÃO	INTENSÃO
O presidente do Brasil	Lula	Líder máximo do Brasil

Quadro 16

A frase “Presidente do Brasil” tem como extensão um referente individual (Lula), mas a sua intensão, noutros termos, o seu alcance, envolve o conceito de algo como “Líder máximo do Brasil”, ou algo como “Sindicalista eleito nas últimas eleições Presidente do Brasil”, ou “O marido da primeira-dama Marisa” etc. Mas, ainda assim, estamos diante de um impasse: a distinção entre extensão e intensão traduz muito pouco do significado de palavras e frases, não resolve o problema do sentido, nem tampouco se relaciona diretamente ao problema da checagem. Mas vamos argumentar ao contrário aqui. E nos impõe uma questão, digamos, mais de fundo: qual é a natureza do sentido inerente ou da intensão no jogo semântico com as palavras e frases de uma língua e suas relações com a checagem?

Uma saída é compreendermos que o sentido (a intensão) corresponda a uma imagem mental do referente em jogo, ou uma certa visão ideológica do referente. A extensão “Lula”

pode evocar coisas neutras como “Líder máximo do Brasil”, ou coisas mais engajadas como “Capa-preta do mensalão” (Folha de São Paulo), ou emotivas como “Presidente brasileiro mais chorão da história” (The Guardian) ou ainda “Presidente cachaceiro” (The New York Times).

O mais surpreende desse tipo de concepção sobre o significado relacionado a checagem é que ele pode levar a mente humana a conceber coisas inimagináveis no mundo real. E traz para a teoria da referência (teoria que procura explicar como as coisas recebem o nome que têm) um avanço e uma sofisticação na medida em que permite a possibilidade de que imagens mentais como a de um dragão, de um unicórnio ou do Papai Noel, entidades que não existem no mundo real, passem a dele fazer parte. Ou ainda conceber entidades que não existiam no mundo real passem a existir, como a bomba atômica, o Viagra, e a Internet.

Infelizmente, essa idéia – a imagem mental que materializa o inexistente – encontra sérias dificuldades em algum nível. Por um lado, é difícil conceber a imagem mental de palavras como “Oxigênio”, “Átomo”, “Liberdade” ou “saudade”. Mas por outro lado, coisas como “cachorro”, “cigarro” ou “disco” são perfeitamente concebíveis mentalmente, e mais, nos ajudariam a incluir “boxer” e “vira-lata” no grupo dos cachorros, e excluir “lobo” e “coelho” desse grupo; e o mesmo pode ser feito com “cigarro” e “disco”, ou seja, incluí-los em alguma intensão de algum nível ou excluí-los também.

Tentativas de “colar” as coisas aos nomes mentalmente são feitas indefinidamente e constantemente pelos falantes de uma língua. Quando um objeto novo surge no mundo real, os falantes são obrigados a nomeá-lo, pois, para ter acesso, e possuir um objeto, é necessária a intermediação da linguagem em suas relações de extensão e intensão com o objeto a ser nomeado.

Os falantes se utilizam de estratégias as mais elementares, como por exemplo, observar um traço – uma característica – peculiar de um objeto e ascendê-la como possibilidade de nomear o objeto todo. Sequências como “o dente do pente”, “o dente de alho” “o dente de leite” gozam de algo em comum: há uma identidade entre essas três coisas, por isso a palavra “dente” não está gratuita nas sequências acima. Outras vezes, os falantes precisam tomar de empréstimo de línguas estrangeiras nomes para nomear as coisas, é o caso de email e pen-drive.

Mas, de uma coisa temos: os nomes não surgem do nada, não são criados por geração espontânea. Eles geralmente brotam de nomes que os falantes já ouviram antes. É como se as coisas novas que vemos estivessem circunscritas em nossa mente antes de as nomearmos, vindas de tempos imemoriais. Ou seja, a checagem das entradas na mente humana aciona

entradas e padrões já conhecidos. Uma outra questão que nos diz respeito diretamente aqui é como a checagem pode ser visualizada no cérebro. Próximas argumentações a ser feitas.

Vamos imaginar a possibilidade de abrir a cabeça de um ser humano vivo, esquadrinhá-la e observar o cérebro durante atividades as quais os seres humanos estivessem desenvolvendo, tais como ler, escrever, assistir um filme ou mesmo falando. O que nós veríamos ali? A resposta é... nada!

Para um observador de fora, o trabalho do cérebro não demonstra evidência nenhuma de sua atividade. Estamos diante de um problema para o campo da neurolinguística, que exige o uso de técnicas investigativas especiais para descobrir os segredos de onde e como a língua é processada no cérebro. Além disso, essas técnicas especiais esbarrariam em exigências éticas de pesquisas tendo seres humanos vivos como objeto. Enquanto outros neurocientistas não encontram lá muitas resistências éticas para a manipulação em laboratório de animais, mesmo ainda vivos, como objeto de suas pesquisas, essa opção não está disponível para os neurolinguistas.

Até recentemente, a única forma de estudar o cérebro era através de estudos feitos por autópsias. Essa técnica era executada nos óbitos de pacientes admitidos em hospitais, geralmente com um quadro de desordem neurológica. Observações cuidadosas eram feitas desde o comportamento do paciente até o seu óbito. O cérebro era examinado para determinar quais áreas foram afetadas. Ao comparar a área afetada do cérebro e o tipo de desordem de que o paciente era acometido quando ainda vivo, os neurologistas desenvolviam teorias sobre o papel da área afetada comparando-a com óbitos de pacientes que não pereceram da desordem quando ainda vivos, e que morreram por outras causas.

Deixando de lado objetos, metodologias e resultados da área da neurolinguística, e nos concentrando em conjecturas sobre como o cérebro funciona na checagem das entradas lexicais, e deixando de lado possibilidades investigativas cujos métodos podem incluir técnicas avançadas como ressonância magnética, esquadrinhamento por computadores, dentre outras, estabelecemos hipóteses sobre o delineamento do cérebro observando os dados da língua no mundo real de fala, ou seja, a língua usada pelos falantes depois que passou da mente (o cérebro em atividade) e saiu da boca das pessoas para contextos sociais.

Tomando os dados do mundo real – frases que os falantes falam para outros falantes – como objeto de pesquisa, hipotetizamos que os falantes fazem movimentos na mente para

formular as frases que dizem para os outros. Como uma hipótese com um nível de abstração como esse pode ser plausível? É o que vamos discutir nos parágrafos abaixo.

Vamos levar em consideração uma categoria (classe) gramatical bem conhecida pelos estudantes de língua portuguesa: o objeto (direto e indireto) dos verbos. Observemos estas sentenças:

- a) Joana viu o rapaz na feira ontem
- b) O rapaz, Joana viu na feira ontem

Veja que o grupo de palavras [o rapaz] tem algo a ver com o verbo [viu], tanto na sentença (a) quanto na sentença (b), ou seja, Joana viu alguma coisa, e foi [o rapaz] que Joana viu. Veja também que o grupo de palavras [o rapaz] aparece logo depois do verbo [viu] na sentença (a), e na periferia à esquerda, ou seja, logo no começo da sentença, em (b). E mais, temos plena consciência de que o significado de [o rapaz] em (a) e em (b) permanece ligeiramente o mesmo, ou seja, permanece tanto em uma sentença quanto em outra portando a mesma informação: “o” alguém que Joana viu na feira ontem.

Partindo de evidências como essas, conjecturamos que as entradas fazem movimentos na mente humana, e que esses movimentos são propriamente movimentos *da* mente humana. As entradas, portanto, nascem em um lugar mas se deslocam para um outro lugar. Ou seja, novamente, o grupo de palavras [o rapaz] nasce de junto do verbo, por ser-lhe seu complemento, por estar diretamente associado ao verbo, mas se desloca para um outro lugar da sentença, como vimos, para o começo da sentença.

Mas por que ou para que existem movimentos na mente humana, tendo a checagem como o fenômeno recortado para demonstrar esses movimentos? Poderemos, rapidamente, deduzir duas coisas: primeira, a composição maleável, moldável do cérebro, que se adapta a circunstâncias que lhe são impostas pela peculiaridade da própria dinâmica humana de ser; segunda, agora, do ponto de vista dos objetivos de fala: os movimentos existem para simular determinadas intenções de fala dos falantes. Veja que o grupo de palavras [o rapaz] que aparece na periferia à esquerda da sentença (b) permanece a mesma categoria, ainda diz alguma coisa sobre o verbo [viu], mas muda sua dimensão semântica. Quando os falantes querem dar relevância a determinados grupos de palavras, uma das estratégias que eles usam é colocar esse grupo de palavras na frente de uma frase, como foi feito com [o rapaz] na sentença (b). Seu significado, portanto, muda ligeiramente quando esse grupo aparece depois do verbo [viu] em (a). Nem precisamos, na verdade, abrir a cabeça de uma pessoa para saber como o seu cérebro funciona: seu jeito de usar a sua língua escancara isso para nós.



Contamos essa passagem para chegarmos a uma conclusão plausível para a checagem: exclusões e padrões, acionados pelo léxico e pela morfologia, assim como a própria forma de ser do cérebro, e os movimentos da mente ajudam os falantes a fazer a checagem das entradas em suas mentes.

**RESUMO:** Morfemas e palavras não conseguem dar conta dos mecanismos de funcionamento das entradas. Propomos o abandono dessas noções, e, em seu lugar, devemos assumir a noção de entrada lexical, pois os falantes atuam em sua língua disponibilizando entradas (palavras e morfemas) de maneira assemelhada: por intermédio de uma leitura semântica. Assumindo esses termos como entradas lexicais, os falantes estão aptos a deslocar as entradas em direção a um sentença, depois de checá-las em suas mentes.

## **9 CONSIDERAÇÕES FINAIS**

Ao longo do trabalho, lançamos algumas conclusões provisórias, que agora serão retomadas e rapidamente comentadas. Chamamos de provisórias pois as discussões posteriores a elas por vezes reformulavam seu teor.

Conclusão provisória 1: há constituintes semânticos menores e maiores de que uma unidade lexical, mas uma unidade lexical não pode prescindir de (por ausência e por presença): uma constituência semântica, em termos de um constituinte semântico; investidas gramaticais, em termos de adequação sintagmática na estrutura de uma língua; e, um tamanho: não pode ser menor de que uma palavra, como também não pode ser maior que uma

palavra – no caso, ser uma sentença. Nesses termos, estabelecemos uma entrada lexical como uma unidade semântica acessada pelos falantes para a construção de sentenças.

Conclusão provisória 2: As entradas lexicais, construído hipotético submetido a verificações em termos mentalistas e em termos públicos (visualizados em dicionários como o Aurélio e o Houaiss), se associam a possibilidades de concatenação do léxico para fins cognitivos e funcionais. Assim que passam por um processo de checagem na mente humana, as entradas lexicais estão disponíveis para uso.

Conclusão provisória 3: em termos de domínio de regras gramaticais podemos afirmar que nossos dois falantes hipotéticos dominam as mesmas regras; seria impossível que eles não as dominassem, pois, além de estarem se comunicando por meio de uma língua, eles se comunicariam por meio de um código similar, a língua portuguesa. Nesses termos, para que haja a comunicação efetiva, os falantes terão que dominar ao menos um traço semântico de uma entrada lexical comum para eles.

Conclusão provisória 4: os verbos, do ponto de vista estrutural e relacional, estão no comando semântico-relacional. Nesses termos, uma entrada lexical verbo é o centro do domínio do léxico.

Conclusão provisória 5: é preciso colocar o traço semântico [+agente] sob rasura quando é pressuposto como argumento de um verbo, pois há itens que, mesmo não sendo argumentos de um verbo que demanda [+agente], mesmo assim continuam sendo [+agente]. Problematizamos o traço semântico agente como relevante para a discussão de entrada lexical.

Conclusão provisória 6: distribuição em paradigmas, ou seja, a ação da morfologia, é acionada pelos falantes na checagem das entradas em nominais ou verbais.

Conclusão provisória 7: o deslocamento do léxico para a estrutura completa o percurso das entradas.

Os estudos sobre aquisição da linguagem empreendidos a partir da aventura científica na investigação trazida pela hipótese inatista em contraponto às ciências do comportamento de base behaviorista se estabelecem como um ponto de inflexão da ciência linguística do século XX.

A hipótese inatista centra suas forças teóricas em como as crianças adquirem a gramática de sua língua nativa, basicamente na investigação das etapas de maturação de um órgão humano herdado pelo processo de evolução que concebemos por abstração – a sintaxe. Como há um período crítico até que a sintaxe seja plenamente de domínio das crianças, propomos, ao longo de nossas discussões, que entradas lexicais são o primeiro contato das crianças com sua língua nativa, em outros termos, propomos que antes da sintaxe as crianças

de fato adquirem entradas lexicais, e a partir delas, a sintaxe começa sua aventura pela mente das crianças até que a gramática de sua língua nativa esteja plenamente estabelecida.

Dois sistemas modulares da mente humana entram em jogo na habilidade de reconhecimento, assimilação e domínio de entradas lexicais: a morfologia e o léxico. A aquisição de novas entradas, atividade contínua e permanente da mente humana desde o momento em que se dispara o processo de estocagem das entradas, pode ser vista como um trabalho linguístico de apelo duplo: as entradas são conseguidas de forma irregular – uma a uma; assim como podem ser conseguidas de forma regular – o acionamento de padrões.

Temos isso como hipótese forte de nosso trabalho na medida em que, por exemplo, em língua portuguesa, o jogo com os padrões, a cargo da morfologia, organiza as entradas na mente humana.

Num dicionário público, como o Aurélio, por exemplo, raramente existem entradas estocadas no singular e no plural. Uma entrada como [jornalista] aparece uma única vez num dicionário público – não há necessidade de grafá-la no plural, pois, para isso, os falantes acionam o padrão do plural em português (ausência X presença de um [s]). Em contato com novas entradas, o jogo com o padrão, novamente, herdado da morfologia, é acionado pelos falantes para depurá-la semântica e estruturalmente. Nesses termos, como demonstramos, uma entrada supostamente nova e desconhecida como [malufar] não terá dificuldades em ser estocada pelos falantes do português, pois a morfologia entra e faz o seu trabalho.

Já a estocagem das entradas irregulares, para as quais o padrão não atua ou atua de forma secundária, entra em cena o léxico. Por exemplo, o gênero em língua portuguesa só se beneficia do padrão em apenas 4,5% das entradas, em palavras como o par menino/menina; o restante tem que ser estocado item a item, já que os falantes precisam estocar o gênero de uma entrada como [tribo], já marcada no léxico como de gênero feminino, levando em consideração uma irregularidade de que o padrão não dá conta.

Temos como hipótese que, na mente humana, há um dicionário similar ao dicionário público, e a estocagem das entradas na mente é feita também de forma similar, concorrendo os dois sistemas: o léxico e a morfologia.

Como as entradas viajam na mente para fora da mente foi alvo de escrutínio do ponto de vista semântico, nesses termos, há uma espécie de escolha das entradas na mente antes do uso. A escolha se faz por intermédio de projeções semânticas, para as quais, concorrem modulação de sentidos e o trabalho de deslizamento semântico já explicitados. Temos como

hipótese que, enquanto para o uso, os rituais sociais de fala são relevantes, do ponto de vista cognitivo, na mente humana, portanto, as projeções semânticas e os ataques gramaticais são essenciais na escolha das entradas.

As entradas lexicais atuam na estrutura em termos composicionais. Isso significa que há uma concatenação, uma espécie de fricção semântica, para que o sentido global de uma sentença seja elaborado pelos falantes. Se não fosse isso, como poderíamos entender que uma entrada como [cavalo] se relaciona de maneira mais próxima a [galopa], e de maneira mais distante com [fala]? Relações composicionais e o trabalho de fricção semântica ajudam os falantes a salvarem sintagmas do tipo [cachorro falante], aparentemente estranhos mas composicionalmente viáveis do ponto de vista semântico.

As entradas em outro foco desta tese foram estabelecidas em variados níveis da língua por se estabelecerem como unidades semânticas. Suas propriedades semânticas se espalham por níveis que vão de átomos linguísticos a palavras, estendendo-se a metáforas. Afixos, palavras e metáforas são organizadas na mente como entradas lexicais pois são percebidas na mente como uma unidade. Se uma palavra é uma unidade na estrutura sentencial, um afixo é uma unidade no interior de uma palavra, assim como uma metáfora é uma unidade do ponto de vista discursivo.

Para termos como hipótese que uma entrada se organiza e é estocada na mente, ou seja, existe na mente como entrada, fizemos a desconstrução de como uma entrada é registrada num dicionário. Vimos que, num dicionário, uma entrada é caracterizada do ponto de vista formal, funcional e semântico.

Em termos formais, um verbete de dicionário apresenta sua forma gráfica, fonológica e morfológica, apresentando sua estrutura. Em termos funcionais, entra em jogo a função gramatical, em que é disponibilizada para o leitor informação distributiva e categorial, assim como é também disponibilizada sua classe. Em termos semânticos, fixa-se o seu sentido. Em muitos dicionários disponibilizam-se também informações discursivas, quando a entrada é vista do ponto de vista familiar etc. Na mente humana, essas informações estão também disponíveis para os falantes, ao mesmo tempo. Na escolha de um item que vai da mente para fora da mente, os falantes atuam através de ataques gramaticais, apagando informações não relevantes, e açando as informações que são necessárias para os interesses dos falantes, como vimos quando fixamos o uso de uma entrada ambígua como [banco].

A aventura linguística humana na Terra começa com uma simples Entrada Lexical.

## 10 REFERÊNCIAS BIBLIOGRÁFICAS

- ARONOFF, Mark & ANSHEN, Frank. Morphology and the Lexicon: Lexicalization and Productivity. In: *The Handbook of Morphology*. Oxford-UK: Blackwell Publishing, 1998, 2001.
- ARONOFF, Mark. *Word in Generative Grammar*. Massachusetts: The MIT Press, 1976.
- AUSTIN, J. L. *How to do things with words*. Oxford: Oxford University Press, 1962.
- BASÍLIO, Margarida. *Estruturas Lexicais do Português: uma abordagem gerativa*. Petrópolis: Vozes, 1980.
- BIDERMAN, Maria Teresa C. *Teoria Linguística (teoria lexical e lingüística computacional) – leitura crítica*. São Paulo: Martins Fontes, 1996, 2001.
- BLOOMFIELD, Leonard. *Language*. New York: Holt, Rinehart and Winston, 1933.
- BRÉAL, Michel. *Essai de sémantique*. Paris: Hachette, 1897.
- BYBEE, J. & SLOBIN, D. I. *Morphological classes as natural categories*. Oxford: Oxford University Press, 1982.

- CÂMARA JR., Joaquim Mattoso. *Estrutura da Língua Portuguesa*. Petrópolis: Vozes, 1970.
- CANÇADO, Márcia. *Manual de Semântica*. Belo Horizonte: Editora da UFMG, 1995.
- CARVALHO, José Cândido. *O Coronel e o Lobisomen*. São Paulo: Record, 1964.
- CHAFE, Wallace L. *Significado e Estrutura Linguística*. Tradução de: Maria Helena de Moura Neves, Odette Gertrudes Luiza Altmann de Souza Campos e Sônia Veasey Rodrigues. São Paulo: Livros Técnicos e Científicos Editora, 1979.
- CHIERCHIA, Gennaro. *Semântica*. Tradução de: Luis Arthur Pagani, Lígia Negri e Rodolfo Ilari. São Paulo: Editora da UNICAMP-EDUEL, 2001.
- CHOMSKY, Noam. *Aspects of the theory of Syntax*. Cambridge: MIT Press, 1965.
- \_\_\_\_\_. *Language and Mind*. New York: Hancourt Brace Jovanovich, 1972.
- \_\_\_\_\_. *O léxico*. In: *Palavras Impossíveis: Língua e Cotidiano*. Tradução: Braulino Pereira de Santana. São Paulo: Nelpa, 2001, 2009.
- \_\_\_\_\_. *Remarks on Nominalization*. In: R. A. Jacobs and P. S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, Ginn, Waltham, Mass, 1970.
- \_\_\_\_\_. *The Minimalist Program*. Cambridge: MIT Press, 1995.
- COSERIU, Eugenio. *Les Procédés sémantiques dans la formation des mots*. In: *Cahiers Feridnand de Saussure*, Madrid: Editorial Gredos, 1968.
- CRUSE, D.A. *Lexical Semantics*. Tradução livre: Braulino Pereira de Santana. Cambridge: Cambridge University Press, 1992.
- CUNHA, Celso & CINTRA, Lindley. *Nova Gramática do Português Contemporâneo*. 2ª Edição. São Paulo: Nova Fronteira, 1985, 1987.
- Dicionário de Francês: Francês/português, Português/Francês. 41ª Edição. São Paulo: Editora Globo, 1999.
- Dicionário Houaiss da Língua Portuguesa. São Paulo: Objetiva, 2001.
- FILMORE, Charles J. Em favor do caso. In: *A Semântica na Linguística Moderna: o Léxico*. Tradução de: Lúcia Maria Pinheiro Lobato. São Paulo: Francisco Alves, 1972, 1977.
- FODOR, Jerry A. Fodor & KATZ, Jerrold J. Estrutura de uma teoria semântica. In: *A Semântica na Linguística Moderna: o Léxico*. Tradução de: Lúcia Maria Pinheiro Lobato. São Paulo: Francisco Alves, 1972, 1977.
- HALLE, M. & MARANTZ, A. *Distributed morphology and the pieces of inflection*. In Halle and Keyser Eds, 1992, 1993.
- HEIDI, Harley. *English Words: A Linguistic Introduction*. Cambridge: Blackwell Publishing, 2005.
- HOEY, S. *Semantics*. London: Longman, 2001.

- HOFFMAN, C. *An Introduction to Bilingualism*. London: Longman, 1982.
- HUMBOLDT, Alexander von. *On human language*. Cambridge: Cambridge University Press, 1994 [1817].
- JAKENDOFF, Ray. *Morphological and semantic regularities in the lexicon*. *Language*, v. 51, n. 3, 1976
- JENSEN, J. T. & STONG-JENSEN, M. *Morphology is in the lexicon*. *Linguistic Inquiry*. Cambridge: Cambridge University Press, 1984.
- KATAMBA, Francis. *Morphology*. New York: St. Martin's Press, 1992.
- KEMPSON, Ruth. *Semântica*. Tradução: Rita de Cássia Machado. Martins Fontes, 1989.
- LAUDANNA, A. & BURANI, C. *Address mechanisms to decomposed lexical entries*. New York: *Linguistics Journal*, 1985.
- LYONS, J. *Introdução à Teoria Linguística*. Tradução: Rosa Virgínia de Mattos e Silva. São Paulo: Cultrix, 1963, 1982.
- MILLER, L. *Knowledge of Meaning*. Cambridge: MIT Press, 1982.
- MIOTO, Carlos et alli. *Manual de sintaxe*. Florianópolis: Insular, 2002.
- Novo Dicionário Aurélio da Língua Portuguesa. 3ª Edição, Revista e Atualizada. São Paulo: Editora Positivo, 2004.
- PALLETIER, M. & SCHUBERT, L. *An Introduction to Semantic Theory*. Cambridge: MIT Press, 1989.
- PERINI, Mario A. *O sintagma nominal do português: estrutura, significado e função*. *Revista de Estudos da Linguagem*. Belo Horizonte: special number, 1996, 1997.
- PINKER, Steven & PRINCE, A. *On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition*. Berkeley: Berkeley Linguistics Society, 1991.
- PINKER, Steven. *O Instinto da Linguagem: como a mente cria a linguagem*. São Paulo: Martins Fontes, 1994, 2002.
- PULCINELI, Eny. *Linguagem e funcionamento*. São Paulo: Martins Fontes, 2003.
- RADFORD, Andrew. *Minimalist Syntax: exploring the structure of English*. Cambridge-UK: Cambridge University Press, 2004.
- RAPOSO, Eduardo. *Teoria da Gramática: a Faculdade da Linguagem*. Lisboa: Editorial Caminho, 1990.

- RIO-TORTO, Graça Maria. *Formação de palavras em português. Aspectos da construção de avaliativos*. Dissertação de Doutorado em Linguística Portuguesa apresentada à Univ. de Coimbra. Coimbra: 1993, 1998.
- ROCHA, Luís Carlos da Silva. *Estruturas Morfológicas do Português*. Belo Horizonte: Editora da UFMG, 1998.
- ROUSSEAU, Jean-Jacques. *Ensaio sobre a origem das línguas*. In: Os Pensadores. Tradução de: Lourdes Santos Machado. São Paulo: Abril, 1982 [1750].
- SANDMANN, Antônio. *Morfologia Lexical*. São Paulo: Contexto, 1992.
- SANTANA, Bráulino Pereira de. *Gênero linguístico: semântica e função*. Vitória da Conquista: Edições UESB, 2007.
- SAUSSURE, Ferdinand. *Cours de Linguistique Generale*. Paris: Hauchette, 1928.
- SCALISE, Sergio. *Generative Morphology*. Dordrecht: Foris, 1986.
- SILVA, Dionísio. *Semântica*. São Paulo: Vozes, 1996.