

REFERÊNCIA:

LEFFA, Vilson J. A resolução da ambigüidade lexical sem apoio do conhecimento de mundo. *Intercâmbio*. São Paulo, PUC : v. 6, Parte I, p. 869-889, 1996..

A RESOLUÇÃO DA AMBIGÜIDADE LEXICAL SEM APOIO DO CONHECIMENTO DE MUNDO

Vilson J. LEFFA (Universidade Católica de Pelotas)

ABSTRACT: One of the most serious problems found in machine translation, which may have interesting implications for language teaching, is the resolution of lexical ambiguity when translating text into another language. One of the most frequent resources, used in this case, has been the use of encyclopedic knowledge. The attempts made so far to endow the computer with a representation of world knowledge, however, have resulted in combinatorial explosions, which turn any system unfeasible. This study investigates word sense ambiguity through selective restrictions, using English as the source language and portuguese as the target language. Initially 10 ambiguous pairs were selected from the source language and analyzed against an English language corpus made up of 10,000,000 words of expository text. Six examples for each pair were randomly picked from the corpus, using a concordancer, and given to bilingual subjects, who should tell which words surrounding the ambiguous pair could be used to disambiguate it. The results showed that every ambiguous pair was preceded or followed by words which had some semantic features that disambiguated the pair. The conclusion is that selective restrictions replace with many advantages the use of encyclopedic knowledge to solve lexical ambiguities.

Introdução

O objetivo deste trabalho é analisar o problema da ambigüidade que pode surgir na transposição de um texto de uma língua para outra. Procura-se especificamente abordar o problema a partir dos recursos lingüísticos do texto, sem apoio do conhecimento de mundo. O pressuposto é de que o leitor, antes de acionar os esquemas cognitivos globais que vão orientá-lo na compreensão do texto, usa informações lingüísticas locais, baseadas nas relações semânticas entre as palavras. A hipótese geral é de que essas relações semânticas locais, pelas restrições que elas impõem entre si, já resolvem a ambigüidade lexical, sem necessidade da âncora de um esquema.

Tradicionalmente a ambigüidade tem sido tratada do ponto de vista do

produtor do texto, como algo que deve ser evitado para que a clareza do texto não seja prejudicada. Um exemplo clássico, das gramáticas normativas, é a anáfora com referentes duplos. A frase (1), por exemplo, pode significar, no nível restrito da sentença, que João ia ao cinema com sua própria irmã ou com a irmã de Pedro.

(1) *João disse ao Pedro que ia ao cinema com sua irmã.*

A ambigüidade, nesses termos, como observa Ferreira (1994), não existe para o produtor do texto. Ele tem bem claro em sua mente o que deseja expressar e provavelmente nem se dá conta do significado alternativo. O problema existe para o receptor do texto, para quem teoricamente três soluções são possíveis: (1) a ambigüidade é resolvida por algum mecanismo desambiguador, usando dados que estão além da frase; (2) a ambigüidade não é resolvida, com prejuízo da compreensão do texto; (3) a ambigüidade é resolvida erroneamente, com a adoção do significado não pretendido pelo autor.

Do ponto de vista das gramáticas há duas posições diferentes em relação à ambigüidade. Para a gramática prescritiva, a ambigüidade é simplesmente um fenômeno que deve ser evitado, mesmo que possa ser resolvido por informações do contexto maior. Já na gramática descritiva, a ambigüidade é um fenômeno natural da língua, e cabe à gramática analisá-lo do modo adequado, mostrando, fora de contexto, a estrutura profunda de cada interpretação. Um exemplo clássico (Chomsky, 1965) é a frase em língua inglesa abaixo (número 2), que pode ter duas estruturas profundas. No primeiro caso, “flying planes” é um sintagma nominal e a frase poderá ser traduzida como “Os aviões podem ser perigosos”. No segundo caso, “Flying planes” não é um sintagma mas uma oração, o que produziria em português a frase “Pilotar aviões pode ser perigoso”.

(2) *Flying planes can be dangerous.*

O exemplo de Chomsky, como se vê, leva a estruturas profundas diferentes e pode ser desambiguado sintaticamente, em nível de classe de palavras, bastando que se dê um rótulo diferente à palavra “flying”, que pode ter o rótulo adjetivo ou o rótulo verbo: se for adjetivo, temos uma frase com uma oração; se for verbo, temos duas orações.

A desambiguação baseada apenas em rótulos ou classes de palavras, entretanto, nem sempre é possível. Muitas ambigüidades envolvem a mesma estrutura, tanto superficial como profunda. O exemplo (3) pode significar tanto “a parede está danificada” como “o muro está danificado”; em ambas as interpretações, cada palavra da frase continua pertencendo à mesma classe gramatical.

(3) *The wall is damaged.*

A solução para este tipo de problema está portanto além do nível sintático da sentença. A frase só pode ser desambiguada, quando o leitor, usando — e relacionando entre si — os dados que estão antes ou depois da frase, aciona um determinado esquema mental para adequadamente incorporar esses dados.

Nos casos em que essa ambigüidade não é resolvida de imediato, o leitor pode inicialmente acionar múltiplos esquemas até a seleção do esquema definitivo ou, o que é mais provável, acionar um esquema provisório, que poderá ser depois confirmado — ou rejeitado e substituído por um outro. Como apenas um esquema será eventualmente confirmado, temos aqui um caso de acionamento de esquemas mutuamente exclusivos.

Pode também acontecer de o autor do texto usar a ambigüidade intencionalmente, obrigando o leitor a acionar esquemas concomitantes que acabam sendo todos confirmados, complementando-se. Acontece, por exemplo, quanto o autor brinca com a língua, jogando com os esquemas do leitor, como se pode observar no exemplo (4). O esquema dominante, para a primeira parte da frase, é de que “andar na linha” significa agir corretamente. Este esquema não é rejeitado mas associa-se ao segundo, na parte final da frase, onde “andar na linha” significa também “caminhar pelos trilhos da via férrea”. Quando a ambigüidade é intencional, podemos chamá-la mais corretamente de duplo sentido, como coloca Ferreira (1994).

(4) *Ela sempre andou na linha — até o dia em que foi atropelada pelo trem.*

A ambigüidade pode envolver casos de polissemia, homonímia e metáfora. Há polissemia quando os diferentes sentidos de uma palavra estão relacionados entre si: os sentidos da palavra *bico*, por exemplo, em expressões como *o bico da ave*, *o bico do sapato*, *o bico de gás* compartilham todos do mesmo traço semântico *extremidade aguçada*. Já a homonímia, que em alguns casos pode até envolver palavras de diferentes etimologias, geralmente não apresenta qualquer relação entre os significados: *manga*, como parte do vestuário, nada tem em comum com *manga*, o fruto da mangueira.

A metáfora é vista neste trabalho como a aquisição de um novo sentido para uma mesma palavra. A palavra *braço*, por exemplo, pode adquirir o sentido de *afluente* na expressão *o braço do rio*. Quando isso acontece, e principalmente se a expressão ficar de uso corrente na língua, o que era inicialmente uma metáfora — uma transposição quase consciente de uma palavra para um outro âmbito semântico — passa a ser, na realidade, mais um caso de polissemia. O aspecto que interessa

neste trabalho é o fato de a palavra adquirir um outro significado quando se transfere para um outro âmbito semântico; *braço* como afluente de um rio é diferente de *braço* como membro superior.

A ambigüidade pode ser acidental, intencional ou interlingüística. Ela é acidental quando produzida não intencionalmente; neste caso ela não existe para o produtor do texto mas apenas para o receptor. É intencional quando existe para ambos; neste caso temos o duplo sentido. E finalmente, ela pode ser interlingüística, isto é, quando existe apenas na passagem de uma língua para outra. Este último caso é especificamente o problema abordado neste trabalho.

1 Definição do problema

A ambigüidade interlingüística ocorre quando uma determinada palavra tem mais de uma acepção na língua fonte e cada uma dessas acepções são expressas por palavras diferentes na língua alvo. Pode envolver casos de polissemia (ex.: *child* em inglês traduzido como *criança* ou *filho* em português), homonímia (ex.: *bank* como *banco* ou *margem*) e metáfora (ex.: *leg* em inglês, que basicamente significa *perna* mas que é traduzido como *pé* em *the leg of the table/o pé da mesa*).

A desambiguação na língua alvo pode apresentar diferentes graus de dificuldade. Os casos de polissemia, por envolverem sentidos relacionados, e portanto contextos semelhantes, são provavelmente mais difíceis do que os casos de homonímia, geralmente encontrados em contextos mais diversificados. Em princípio a desambiguação será tanto mais difícil quanto mais sua solução se afastar da estrutura superficial da frase, indo em direção à semântica e ao nível do discurso. A tradução de *hope* em inglês para *esperança* ou *esperar* em português é mais fácil porque a palavra pertence a classes gramaticais diferentes, isto é, a ambigüidade já pode ser resolvida no momento em que se decidir a que classe pertence a palavra. Obviamente, quando a palavra pertence à mesma classe gramatical, esse recurso não pode ser usado, e a dificuldade aumenta porque será necessário incluir informações semânticas.

Parece também que algumas classes gramaticais são mais facilmente desambiguadas do que outras. Verbos e adjetivos, por exemplo, têm mais probabilidade de ser desambiguados dentro da própria frase do que os substantivos. O verbo *work* poderá ser traduzido como *trabalhar* se o sujeito for animado mas será provavelmente *funcionar* se o sujeito for inanimado: compare *The man works/O homem trabalha* com *The machine works/A máquina funciona*. Do mesmo modo, as acepções de um adjetivo ambíguo podem ser determinadas pelos traços semânticos do substantivo que qualifica: *short* será *baixo* qualificando um substantivo com o traço semântico *altura*; *curto* se o substantivo

contiver o traço *comprimento* (Justeson & Katz, 1995).

Com o substantivo, as dificuldades são maiores. O substantivo geralmente desempenha a função não de adjunto mas de núcleo e isso o torna menos dependente das palavras que o cercam. Sua desambiguação depende de traços mais afastados, às vezes até além do texto, encontrados apenas nos esquemas que o tradutor aciona para processar o texto. Uma sentença como *Look at the table* pode significar tanto *Olhe para a mesa*, como *Olhe para o quadro*. Fora de um esquema adequado, não há na frase, qualquer indício para sua desambiguação.

A afirmação, no entanto, de que a ambigüidade fica resolvida quando o leitor aciona o esquema adequado apenas joga o problema para frente sem resolvê-lo na realidade. A questão é justamente saber o que leva o leitor a acionar um esquema e não outro. Quais são os fatores que determinam essa escolha? Onde estão estes fatores? Localizar esses fatores é o objetivo deste trabalho. Como o problema maior parece residir na desambiguação de substantivos, o trabalho se fixará especificamente nesse tipo de ambigüidade.

3 Hipótese

A principal hipótese deste trabalho é que a desambiguação das diferentes acepções de um substantivo pode ser resolvida através de dados detectados no próprio texto. Esses dados podem ser a função sintática do substantivo em questão, a existência ou não no texto de uma palavra que pertença a um determinado campo semântico, o tipo de relação semântica que o substantivo pode manter com outras classes de palavras, etc. Esses dados podem estar no próprio substantivo, precedê-lo na frase, precedê-lo no texto, segui-lo na frase ou no texto. Podem ser apenas dados isolados ou uma combinação desses dados. Podem ter também um número maior ou menor de restrições para que possam ser aproveitados. A hipótese parte da premissa de Fillmore de que uma palavra é conhecida pela companhia com que anda.

3 Metodologia

Partindo do pressuposto de que a compreensão, e conseqüentemente a desambiguação, só é possível quando o leitor aciona um esquema onde possa colocar os dados do texto, este trabalho vai tentar descrever o que acontece antes de o leitor acionar o esquema. Para isso usou-se o seguinte procedimento.

Inicialmente selecionaram-se 10 palavras ambíguas, e escolheram-se para cada palavra 2 acepções. A Tabela 1 mostra a lista dessas

palavras em língua inglesa com suas respectivas acepções em português. Os exemplos foram tirados de um corpus de 10 milhões de palavras de textos expositivos em língua inglesa, usando um concordancer (Scott e Johns, 1992), programa que alinha na tela do computador cada palavra selecionada, mostrando o contexto imediato tanto à esquerda como a direita.

Foram adotados três critérios para a seleção dessas palavras: (1) elas deveriam apresentar, no corpus, pelo menos 10 ocorrências para cada uma das duas acepções escolhidas; (2) os exemplos seriam selecionados de modo totalmente aleatório; (3) a palavra ambígua não deveria fazer parte de uma expressão idiomática (ex.: *periodic table*). Partindo das 10 palavras selecionadas e de suas acepções, elaborou-se uma tarefa desambiguadora, composta de 60 itens, apresentando, portanto, 6 itens para cada uma das 10 palavras. Cada um dos 60 itens, por sua vez, continha um segmento de texto, retirado do corpus, com 20 palavras à esquerda e 20 palavras à direita do termo ambíguo. A Figura 1 mostra 2 desses itens, com o texto, as acepções e percentual de certeza a ser preenchido pelo leitor.

Tabela 1 — Palavras e acepções selecionadas para o tarefa desambiguadora.

Termo em inglês	Acepção 1	Acepção 2
arm	arma	braço
bank	banco	margem
board	conselho	prancha
coach	carruagem	treinador
letter	carta	letra
nail	prego	unha
page	pagina	página
table	mesa	tabela
wall	muro	parede
way	modo	caminho

Foram usados como sujeitos 22 professores de língua inglesa, incluindo colegas do pesquisador, alunos que participavam de um curso de mestrado e professores de dois cursos de especialização em língua inglesa. Embora o grau de proficiência na língua variasse entre esses sujeitos, todos possuíam o nível suficiente para ler e compreender os textos da tarefa.

<p>...as peace commissioners. They had little with which to negotiate, except the vague assertion that Americans should lay down their arms and trust to the good will of Parliament to remove objectionable laws. Even had they possessed meaningful bargaining powers, the...</p>	<p>() braço () arma Certeza _____%</p>
<p>...cord to the muscles. Degeneration of these neurons causes muscle weakness, spasticity, and atrophy, usually starting in the hands and arms and then spreading to other parts of the body. Difficulty with speaking, swallowing, and breathing ensues, and death usually follows...</p>	<p>() braço () arma Certeza _____%</p>

Figura 1 — Exemplo de contextos imediatos para a palavra *arm*

O procedimento para a coleta dos dados consistiu basicamente de quatro etapas: (1) leitura silenciosa do texto de cada item; (2) escolha da acepção correta; (3) determinação do percentual de certeza; (4) marcação no texto das palavras que mais ajudaram na desambiguação do termo. Para essa última etapa, solicitou-se a cada sujeito que marcasse três palavras no texto, escrevendo um número sobre cada uma, ou seja, o número 1 para a palavra que mais ajudou, o número 2 para a palavra que ajudou em segundo lugar, e o número 3 para a palavra que ajudou em terceiro lugar.

As sessões, com pequenos grupos de professores, foram conduzidas pelo próprio pesquisador. Cada sessão era precedida de um período de treinamento, para que os sujeitos se familiarizassem com as tarefas solicitadas. Não houve pressão de tempo para o término da tarefa, que tomou cerca de 60 minutos, embora em algumas sessões alguns professores obrigaram-se a responder mais rapidamente as últimas questões para não se atrasarem em relação aos que já tinha terminado. Todos os encontros, sempre em pequenos grupos, desenrolaram-se de modo informal e descontraído, empenhando-se todos em executar a tarefa da melhor maneira possível.

4 Resultados e discussão

A Tabela 2 resume os dados obtidos na tarefa desambiguadora. Na maior parte dos casos (77%), a ambigüidade foi desfeita usando um dado desambiguador que se encontrava à esquerda do termo alvo. A distância média entre o dado desambiguador e o termo ambíguo foi de 4,8 palavras, tanto à esquerda como à direita. Essa distância, no entanto, refere-se ao dado que foi considerado o mais útil. Levando em consideração o dado que está mais próximo, a distância diminui para 2,1 palavras.

A Figura 2, com dados retirados aleatoriamente do corpus para a palavra *board*, mostra a proximidade desses dados. Dos seis exemplos, quatro têm o termo desambiguador junto à palavra ambígua.

Usando o princípio das seleções restritivas, podemos argumentar que as palavras ambíguas possuem traços semânticos mutuamente exclusivos, de modo que quando um está presente o outro traço tem que estar necessariamente ausente. A presença ou ausência de cada traço é determinada pelo contexto, entendendo-se aqui como contexto palavras na vizinhança que possuam um ou outro dos traços. Supondo, por exemplo, que a palavra inglesa *board* possa ter, entre outros traços semânticos, o traço *c*, relacionado a *conselho* e o traço *p*, relacionado a *prancha*, ela vai significar *conselho*, se tiver à esquerda ou à direita, palavras que contenham o traço *c* — no exemplo dado seriam *chairman*, *advisory*, *elected* e *dealing*; vai significar *prancha* se, por outro lado, tiver à esquerda ou à direita, palavras com o traço *p* — no caso as palavras *wooden* e *box*.

TABELA 2 — Dados da tarefa desambiguadora

Contexto anterior:	77%
Contexto posterior:	23%
Distância do termo mais útil:	4,8 palavras
Distância do termo mais próximo:	2,1 palavras
Percentual de acertos:	97%
Percentual de certeza:	95%

Texto ambíguo em inglês	Tradução para o português
... <i>chairman</i> of the boardpresidente do conselho ...
..... <i>advisory board</i> conselho consultivo...
..... <i>wooden boards</i> pranchas de madeira...
..... <i>elected boards</i> conselhos eleitos...
..... boards <i>dealing</i> conselhos que tratam...
..... boards or <i>box</i> pranchas ou caixa...

FIGURA 2 — Contexto mínimo de desambiguação para *board* como *prancha* ou *conselho*
(Palavra ambígua em **negrito**, dado desambiguador em *itálico*)

Palavras com traços mutuamente exclusivos não podem ser unificadas — isto é formar um mesmo sintagma (ex.: *presidente da prancha* em vez *presidente do conselho*, *conselhos de madeira* em vez de *pranchas de madeira*) ou mesmo unidades maiores (ex.: *pranchas que tratam* em vez que *conselhos que tratam*). Há uma espécie de valência entre as palavras. Quando determinados traços são compartilhados, pode-se dizer que as palavras se atraem mutuamente, estabelecem uma ligação entre si, constituindo-se em unidades maiores; por haver um compartilhamento de traços entre *conselho* e *consultivo* é possível obter-se *conselho consultivo*. Por outro lado, quando os traços não são compartilhados, as palavras se repelem mutuamente; juntando-se *prancha* com *consultiva* não se obtém *prancha consultiva*. São raros os casos em que o termo desambiguador não aparece no texto, entendendo-se por termo desambiguador aquela palavra que possui o traço semântico que permite ou não a unificação com o termo ambíguo. Dos 60 itens do teste, o único que não apresenta esse termo desambiguador de modo claro é o item abaixo:

... *virginica*, is a solitary bee that tunnels into wood and lays her eggs on a bed of pollen. She builds **walls** of sawdust and saliva between each egg. The female halictid mining bee, *A. metallica*, is a solitary bee. She shares ...

Mesmo assim, poderia-se argumentar que a palavra *egg*, usada duas vezes no segmento, poderia ser o elemento desambiguador. Bastaria para isso dar à palavra *egg* o traço semântico de parte do corpo, que é um dos traços que dá a palavra *wall* o sentido de *parede*, conforme mostra a pesquisa feita no corpus (ex.: *abdominal wall*, *artery wall*, *cell wall*, *vessel wall*).

Mesmo que se rejeitasse esse recurso, haveria ainda a hipótese do sentido dominante (Hogaboam & Perfeti, 1975, *apud* Hirst, 1987), pela qual o sentido de uma palavra ambígua, enquanto não rejeitado pelo contexto, é determinado pelo seu significado mais comum. Dada uma palavra como *pen* numa frase de contexto neutro como *I need a pen/ preciso de uma caneta*, é provável que a maioria dos falantes da língua portuguesa a traduzissem simplesmente pelo seu sentido dominante de *caneta*, sem mesmo se darem conta que *pen* também tem o sentido, muito menos comum, de *aprisco*. Este sentido provavelmente só seria percebido em frases de contexto carregado como *They were treated like sheep in a pen/Foram tratados como ovelhas num aprisco* (tradução literal), onde a palavra *sheep* deixa bem claro que *pen* neste caso não

pode ser *caneta*. No corpus usado nesta investigação, por exemplo, a palavra *wall*, no sentido de *parede*, pode ser considerada mais comum. O sentido dominante de uma palavra ambígua pode ser tratado como um valor *default*, isto é, como um sentido que é automaticamente atribuído à palavra — até que seja confirmado, rejeitado ou não resolvido pelo contexto. Neste último caso permanece o valor *default*. O uso do valor *default* é um recurso que não chegou a ser usado nos 60 segmentos extraídos para a elaboração do teste, já que em todos eles a ambigüidade foi resolvida pelo uso das relações semânticas. Notou-se, porém, uma dificuldade maior quando o termo desambiguador na língua fonte transforma-se num termo ambíguo na língua alvo. Isso aconteceu num dos itens do teste, que é o segmento abaixo:

*... 31, 1870, was inventive in many areas. The son of a Massachusetts machine-shop operator, he devised machines for making **nails** and cutting files while still in his teens. In 1819 he adapted a European process for making patent leather and ...*

Tudo neste segmento parece ter conspirado para complicar a solução da ambigüidade. Em primeiro lugar, *nail* é usado na acepção menos comum de *prego*. Em segundo lugar, a palavra *file*, por sua vez também ambígua, é também usada no sentido menos comum de *lima* (instrumento de lixar). Em terceiro lugar, para completar a lista de pistas falsas, a própria palavra *file* no sentido do *lima* associa-se mais a *unha* do que a *prego*. Na realidade o termo que desambigua *nails* é a palavra *making*, usada aqui no sentido de *fabricar*, sendo mais provável que se fabrique pregos do que unhas. No entanto, como em português a palavra *fabricar*, menos ambígua, é sinônima de *fazer*, mais ambígua, gera-se acidentalmente a ambigüidade entre *fabricar pregos* e *fazer as unhas*. Essa ambigüidade acidental parece ter confundido alguns sujeitos, quando na realidade, em inglês, a ambigüidade não existe: o verbo *making*, pelo traço semântico que compartilha com *nail*, deixa bem claro que se trata de pregos e não de unhas.

Rigorosamente, portanto, nos 60 itens do teste, não chegou a haver necessidade de se usar o valor *default*. Isso significa que pelo menos nesses 60 segmentos da amostra, que foram retirados aleatoriamente de um corpus de 10 milhões de palavras, o uso das regras desambiguadoras, baseadas nas restrições seletivas, resolveu todos os casos. O uso do valor *default* ficou apenas como uma reserva de segurança. O uso das regras desambiguadores, usando apenas dados lingüísticos locais, demonstrou um percentual de acertos de 100%, o que é superior ao nível de acertos dos próprios sujeitos, que foi de 97%, e ao nível de certeza em suas respostas, que foi de 95%.

A solução da ambigüidade pelo uso das restrições semânticas está

baseada na probabilidade e oferece sempre um grau de incerteza. Na medida em que se propõe uma abordagem baseada apenas nas relações semânticas entre as palavras, sem o uso de esquemas, essa incerteza deveria aumentar. Não foi, no entanto o que se observou neste trabalho; o acionamento de um esquema não é garantia de acerto, já que o leitor pode acionar um esquema inadequado ou errado. Mais importante do que acionar um esquema é usar os dados corretos do texto. Os resultados deste trabalho parecem indicar que o uso adequado dos dados locais do texto reduz a incerteza de modo mais eficiente do que o uso de esquemas.

A resolução da ambigüidade é uma tarefa difícil, até para leitores proficientes. No texto abaixo, por exemplo, não houve concordância entre os sujeitos quanto ao sentido da palavra *arm*, às vezes traduzido como *arma*, outras vezes como *braço*.

*...and regimental artillery. As the standing-army principle that began about 1500 came into general use, artillery became an organized **arm** of the military. Louis XIV of France raised a regiment of artillerymen in 1671 and established schools of artillery instruction. (...)*

Em entrevista informal com os sujeitos, após o teste, ouviu-se dois tipos de explicação: (1) a infantaria, para uns, é uma arma do exército; (2) para outros é um braço, no sentido de divisão. Na realidade, se tivessem seguido apenas as pistas locais do texto, usando os traços semânticos comuns que a palavra *arm* compartilha com as palavras *artillery* e *military*, e traduzido *arm* como *arma*, teriam acertado. Na medida, porém, que ignoraram esses dados locais e se ancoraram num esquema inadequado, acabaram produzindo uma tradução errônea. Se isso acontece com leitores relativamente proficientes como o que foram selecionados no experimento, alertados anteriormente para a ambigüidade da palavra, é de se esperar que a contradição aumente com leitores mais ingênuos e não alertados para o significado menos proeminente.

Conclusão

O objetivo deste estudo foi avaliar a possibilidade de resolução da ambigüidade lexical sem o uso do conhecimento de mundo, entendendo-se por conhecimento de mundo o acionamento de esquemas pelo leitor, espécies de cenários mentais, onde as lacunas e ambigüidades existentes no texto são preenchidas e resolvidas.

Pressupondo que a dificuldade maior de desambiguação está nos substantivos, selecionou-se 10 pares de substantivos ambíguos em inglês, retirando-se aleatoriamente de um corpus de 10 milhões de

palavras 6 exemplos para cada par, elaborando-se um teste de 60 itens, que foi posteriormente aplicado em sujeitos falantes do inglês e do português.

Pressupondo que o leitor aciona algum dado do texto, antes de acionar o esquema necessário para resolver o termo ambíguo, o objetivo principal do teste foi determinar onde no texto poderia ser encontrado esse dado desambiguador. O levantamento feito com as palavras ambíguas mostrou que na maioria dos casos o dado desambiguador estava à esquerda do termo ambíguo a uma distância média de 2 palavras para o termo mais próximo e de 5 palavras para o termo considerado mais útil na desambiguação. O uso de relações semânticas entre a palavra ambígua e as palavras desambiguadoras, solucionou 100% dos 60 itens do teste. A estratégia do valor *default*, ou seja, o uso do sentido mais comum da palavra quando não é possível determinar as relações semânticas, não chegou a ser usado, permanecendo apenas como uma reserva de segurança.

A principal conclusão deste trabalho é de que não há necessidade de se acionar um esquema para resolver os problemas de ambigüidade lexical; basta detectar a presença de determinados traços que sejam compartilhados tanto pelo termo desambiguador como pelo termo ambíguo.

Usou-se neste estudo textos autênticos, de caráter expositivo.

Teoricamente é sempre possível fabricar um texto, procedimento muito comum na literatura da teoria de esquemas, e mostrar que sua compreensão só é possível com o acionamento do esquema adequado. O texto abaixo, que só pode ser compreendido na medida em que o leitor acionar o esquema “como fazer o vizinho se mudar”, é um exemplo típico:

Primeiro Joelma soltou uma gaiola de ratos, que tinha conseguido de um amigo maluco. O plano, porém, não deu certo porque o gato acabou com eles. Depois ela deu uma festa para uma turma de adolescentes, e ficou encantada quando todos vieram de moto. Infelizmente não conseguiu a sirene, que pretendia acoplar ao jogo de luzes. Além do mais, o conjunto de som não era suficientemente alto. O pó de mico podia dar certo, mas ela não tinha um ventilador com a potência necessária. As chamadas obscenas pelo telefone deram-lhe alguma esperança, mas aí trocaram o número. Pensou em chamar um vendedor ambulante, mas decidiu armar um longo varal de roupas. O que acabou resolvendo foi a instalação de um cartaz luminoso do outro lado da rua, que ficava piscando todas as noites. Quando o anúncio saiu na seção dos classificados, ela anotou o telefone feliz da vida (Bransford et al., 1984, p. 32).

Trata-se de um texto totalmente fabricado, feito com o cuidado

específico de não fornecer qualquer pista ao leitor. Essa escamoteamento de pistas, porém, não é um procedimento normal no mundo dos textos autênticos. Aí as pistas parecem predominar, e podem até dispensar a necessidade do uso de esquemas — pelo menos na resolução da ambigüidade lexical.

REFERÊNCIAS BIBLIOGRÁFICAS

- BRANSFORD, J. D., STEIN, B. S., SHELTON, T. (1984) *Learning from the perspective of the comprehender*. In: ALDERSON, J. C., URQUHART, A. A., orgs. **Reading in a foreign language**. London: Longman.
- CHOMSKY, N. (1965) **Aspects of the theory of syntax**. Cambridge, Mass.: MIT Press.
- FERREIRA, M. C. L. (1994) **A resistência da língua nos limites da sintaxe e do discurso: da ambigüidade ao equívoco**. Campinas: UNICAMP. [Tese de doutorado]
- HIRST, G. (1987) **Semantic interpretation and the resolution of ambiguity**. Cambridge, University Press.
- HOGABOAM, T. W.; PERFETTI, C. A. (1975) *Lexical ambiguity and sentence comprehension*. **Journal of Verbal Behavior**, 16(3): 265-274.
- JUSTESON, J. S. & KATZ, S. M. (1995) *Principled disambiguation: Discriminating adjective senses with modified nouns*. **Computational Linguistics**, 21(1):1-27.
- SCOTT, M. & JOHNS, T. (1992) **Microconcord**. Oxford: University Press. (Oxford English Software).