

Nota: A versão publicada deste texto está sem as figuras e, por isso, mais resumida)

**REFERÊNCIA:** LEFFA, Vilson J. A língua portuguesa no novo milênio: a pesquisa como fonte de conhecimento para o ensino. In HENRIQUES, Cláudio Cezar; PEREIRA, Maria Teresa Gonçalves (orgs.) *Língua e transdisciplinaridade: rumos, conexões, sentidos*. São Paulo: Contexto, 2002, p. 25-42.

## A língua portuguesa no novo milênio

### A pesquisa como fonte de conhecimento para o ensino

Vilson J. Leffa (Universidade Católica de Pelotas)

O objetivo desta apresentação é mostrar a importância da pesquisa no ensino da língua portuguesa. Parte-se de dois princípios básicos aqui. O primeiro é que o ensino envolve o conhecimento, o que é óbvio; ensinar é no mínimo transmitir conhecimento para alguém. O segundo, não tão óbvio, é que o professor deve não apenas passar o conhecimento para o aluno, mas também gerar o conhecimento que passa. Esta é a tese, o argumento principal desta apresentação: como gerar o conhecimento que se passa ao aluno.

#### CONHECIMENTO RECEBIDO

O conhecimento pode ser recebido ou construído. O conhecimento recebido é aquele que vem de uma autoridade e já pronto para consumo; tipicamente é aquele que se recebe na escola, nos livros, nas publicações acadêmicas e até nos meios de comunicação de massa. O uso da autoridade é bastante comum na divulgação do conhecimento sobre a língua. As principais gramáticas da língua portuguesa, por exemplo, são conhecidas pelos nomes de seus autores. Temos, assim, a gramática do Bechara, do Celso Cunha, do Napoleão, para citar apenas três. O mesmo ocorre com os dicionários. Os três maiores dicionários da língua portuguesa são conhecidos pelo nome das pessoas que os idealizaram: o Aurélio, o Michaelis e, mais recentemente, o Houaiss. É óbvio que um dicionário do porte de qualquer um desses três não pode ser a obra de uma única pessoa. O conhecimento necessário para produzir um dicionário – hoje – é tão complexo que não pode ser mais o resultado de um trabalho individual; tem que ser um trabalho de equipe. Estamos na época da inteligência coletiva, distribuída não só entre pessoas afinadas por um mesmo objetivo, mas até entre pessoas e máquinas. O Houaiss, para ficar ainda no exemplo do dicionário, não seria viável na década de 60, nem em termos de qualidade de impressão nem em termos de custo. Não haveria, na época, a tecnologia disponível para apresentar um produto que reunisse a qualidade gráfica e científica que se exige hoje de um dicionário, a um preço que possa ser pago, pelo menos, por um segmento razoável da população.

A diferença entre inteligência individual e coletiva pode ser demonstrada através dos diferentes resultados que ela produz. O que se pensa individualmente é mais limitado e produz resultados mais pobres do que aquilo que é pensado coletivamente. Vejamos como exemplo a definição da palavra *sílaba* em duas edições do dicionário Aurélio. Na segunda edição do dicionário, de 1986, há três acepções para sílaba:

**Sílaba** [Do gr. *syllabé*, pelo lat. *syllaba*.] *s. f.* **1.** Som produzido por uma única emissão de voz assinalada por um ápice de abrimento articulatório e tensão muscular, que corresponde, em português, a uma vogal. **2.** Vogal ou reunião de fonemas que se pronunciam numa só emissão de voz. **3.** Fig. e fam. Som articulado. [Cf. *silaba*, do v. *silabar*.] (...) (Ferreira, 1986, p. 1584).

Não sei qual foi o critério usado pelo autor para ordenar as acepções no verbete, mas imagino que deve ter tido a preocupação de colocar em primeiro lugar a acepção mais freqüente, mais comum na língua. Um outro aspecto que salta aos olhos é que não houve qualquer preocupação em definir os contextos, as diferentes áreas de conhecimento, em que pode aparecer cada uma das acepções.

Mas mesmo assim, mesmo que não tenha priorizado a freqüência de uso, mesmo que não tenha especificado a área científica em que pode ser usado o conceito, a definição oferecida é no mínimo incompleta. Imagino que para todos os professores de língua portuguesa, fonoaudiólogos e muitos alunos de letras, definir a sílaba, em primeira opção, como uma vogal, é certamente incorreta: atribui universalidade a um traço, a uma característica que é apenas circunstancial. Exemplificando, seria mais ou menos como afirmar que as mulheres não sabem dirigir automóvel; talvez algumas não saibam, mas não se pode universalizar a partir daí. Há muitas mulheres que sabem; há muitas que dirigem melhor do que alguns homens, etc.

Note-se, no entanto, a mudança que ocorreu na terceira edição do dicionário, publicada em 1999:

**Sílaba** [Do gr. *syllabé*, pelo lat. *syllaba*.] *S. f.* **1.** *E. Ling.* Na cadeia da fala, cada uma das unidades estruturadas com base em seqüências de vogais e consoantes e que se reúnem para constituir as palavras. [Cf. *coda*, *núcleo*, *onset*, *rima*<sup>1</sup>.] **2.** *E. Ling.* A representação dessas unidades na escrita. **3.** *Fig. Fam.* Som articulado. [Cf. *silaba*, do v. *silabar*.] (...) (Ferreira, 1999, p. 1853).

Também aqui são apresentadas três acepções, mas as semelhanças param aí. De imediato, nota-se que cada acepção é contextualizada, sendo que as duas primeiras pertencem à área de Estudos da Linguagem. Pode haver ainda alguns problemas. Pessoalmente, não vejo a necessidade de separar a sílaba na “cadeia de fala” e na escrita. É possível também que a expressão “unidades estruturadas com base em seqüências de vogais e consoantes” não ajude muito a quem vai ao dicionário para saber o que é sílaba; pode ser que seja até totalmente inútil, mas pelo menos, a meu ver, não se dá ao consulente uma informação errada. A remissão que o dicionário faz para as palavras *coda*, *núcleo*, *onset*, *rima* me pareceu, por outro lado, de extrema utilidade – e se um leitor tiver um mínimo de curiosidade e for investigar cada um desses termos, terá a oportunidade não só de

adquirir o conceito correto de sílaba mas também de se atualizar na terminologia adequada. Vejamos os exemplos retirados do dicionário, reproduzidos aqui na acepção referente à área da linguagem:

**Núcleo** Na estrutura da sílaba, a posição obrigatória, ocupada pela vogal.

**Onset** Na estrutura da sílaba, a posição que precede a vogal e que pode estar vazia ou ser ocupada por uma ou duas consoantes.

**Coda** Segmento(s) consonântico(s) que ocupa(m) a parte final da sílaba.

**Rima** Na estrutura da sílaba, unidade formada pelo núcleo e pela coda.

Considerando que o conceito de sílaba deve ter permanecido o mesmo entre 1986 (data de segunda edição do dicionário) e 1999 (data da terceira), essa mudança na definição deve ser atribuída não à evolução da lingüística mas a uma nova maneira de trabalhar. Vejo, nessa nova maneira, a passagem do individual para o coletivo.

Há um poema de Emily Dickinson que, em tradução livre para o português, diz mais ou menos o seguinte:

Para fazer uma pradaria  
é preciso uma abelha e um trevo  
mas se não tiver um trevo  
uma abelha é suficiente

Acho esse poema muito bonito, mas vejo nele um problema. Na minha leitura, ele é uma exaltação ao individualismo como o valor mais elevado da humanidade; uma glorificação do indivíduo que faz e acontece e não precisa da ajuda dos outros. Há no fundo, a idéia romântica do lobo solitário na estepe, totalmente auto-suficiente e dono da planície. Isso até pode ser bonito, mas não é verdade; ninguém é capaz de construir uma pradaria sozinho – como ninguém é capaz de produzir um dicionário sozinho.

Antigamente distribuíam-se tarefas; hoje temos que aprender a distribuir o pensamento. As idéias e a própria inteligência têm que transbordar do individual para o coletivo. Estamos entrando numa sociedade onde o bem maior é o conhecimento, mas, por mais que se defina e se recorte esse conhecimento, por mais especializado que ele seja, ele é ainda grande demais para caber num único cérebro. Ninguém vai ter um cérebro que possa conter uma mente que seja tão desenvolvida que possa reter todo o conhecimento necessário para produzir um dicionário. O conhecimento precisa ser distribuído entre cérebros diferentes. Eu diria até mais, ainda que com medo de ser vaiado: o conhecimento precisa ser distribuído não só entre cérebros humanos mas também incluir a memória das máquinas. Não só a impressão do texto no papel, com o uso de cores diferentes, ilustrações,

etc, é feita por máquinas, mas também processos envolvendo cálculos, comparações, remissões a outras palavras, etc. A complexidade do conhecimento exige que a inteligência deixe de ser um atributo individual, estanque dentro das pessoas, para se tornar uma substância fluida, distribuída coletivamente, por um processo de transbordamento.

Mesmo com toda distribuição e integração de tarefas, mecânicas e cognitivas, que caracteriza a feitura de um dicionário, mesmo assim, ao ser consultado, ele se torna para o leitor, conhecimento recebido, pronto para ser usado. O perigo é encarar esse conhecimento recebido como autoridade inquestionável para resolver qualquer dúvida. Se o Aurélio disse que sílaba é vogal, então está dito e não se discute mais. É aquela tradição do dicionário percebido como “amansa burro”.

O uso da autoridade como fonte de conhecimento pode representar no máximo uma meia verdade. Se eu digo, por exemplo, que língua é um instinto que as pessoas trazem de nascença e cito Chomsky para corroborar o que digo, eu posso estar totalmente enganado ou apenas dizendo uma meia verdade. Posso afirmar também o contrário e dizer que língua não é um instinto de nascença, mas o produto de uma mente social e cito então Vygotsky para apoiar o que afirmo. Por mais autoridade que tenham Chomsky ou Vygotsky, ao usá-los como fonte de conhecimento, estou correndo o risco de ficar apenas com uma verdade provisória, momentaneamente privilegiada por alguns teóricos mas categoricamente negada por outros.

Alguns teóricos poderão ter pensado muito, dito coisas interessantes que gostamos de ler, mas não são necessariamente os donos da verdade. Como disse Diderot, numa frase extremamente espirituosa: a verdade é uma prostituta; dorme com todos e não é fiel a ninguém. As palavras de uma autoridade, incluindo obviamente o que disse Diderot, não devem ser vistas como fonte da verdade mas como hipóteses, verdades provisórias que poderão ou não ser confirmadas. Podem ser o caminho, podem servir de incentivo e até de inspiração na busca do conhecimento, mas não representam, em si, a verdade.

## CONHECIMENTO CONSTRUÍDO

O conhecimento construído é aquele que surge da própria experiência de vida – passando pela reflexão. A combinação da experiência com a reflexão é extremamente importante e pode ser mostrada graficamente através da expressão mostrada na Figura 1, apresentada aqui apenas para realçar visualmente a relação entre as três variáveis.

**Conhecimento construído = Experiência + Reflexão**

Figura 1 – Definição de conhecimento construído

Essas relações não são estáticas mas dinâmicas, tendo não apenas uma dimensão espacial, mas também temporal. Para entender essas relações dinâmicas – como algo que acontece no tempo e não apenas algo que está disposto graficamente – talvez fosse interessante comparar a produção do

conhecimento construído ao processo biológico da ruminação.

Sabemos que o estômago dos ruminantes tem quatro compartimentos: um primeiro – que se chama rúmen, em linguagem mais acadêmica, ou pança, em linguagem mais popular – e que é onde se guarda o alimento ingerido que não foi mastigado. Esse alimento passa por um processo inicial de fermentação, é depois regurgitado, mastigado várias vezes e novamente deglutido, desta vez para o segundo compartimento, o retículo, de onde vai para o terceiro e então para o quarto, onde é finalmente digerido pelo organismo. O alimento em estado bruto é assim trabalhado pela ruminação – tal como a experiência, que para ser assimilada e transformada em conhecimento, tem que ser trabalhada pela reflexão.

Como se vê, é bastante óbvia a relação entre ruminar e refletir. O que pode não ser muito claro – e precisa ser repetido – é que refletir sobre a experiência pode transformá-la em conhecimento. Em outras palavras, quem tem experiências e reflete sobre essas experiências, pode construir conhecimento e se transformar em pesquisador.

A questão que se coloca neste ponto é se o professor pode produzir o conhecimento que “passa” aos alunos, ou se ele é apenas um intermediário, um atravessador do conhecimento.

Para alguns, o professor não tem condições de pesquisar. Pode ter, quando muito, uma “uma atitude investigativa” (Ferreiro, 2001, p. 36). Tentar fazer do professor um pesquisador é populismo.

Ainda, segundo Ferreiro:

Quando se faz investigação-ação com a idéia de que os professores são investigadores, eu digo que não é bem assim. Ser pesquisador é uma profissão como qualquer outra, nem mais, nem menos. (...) todos podemos cantar, mas há alguns que são cantores profissionais, especializados, e ninguém pode transformar-se em cantor por uma decisão coletiva (Ferreiro, 2001, p. 36).

Trago Emilia Ferreiro para o debate da questão ensino/pesquisa para mostrar que não há unanimidade sobre o assunto. Há vozes discordantes, embora a grande maioria seja a favor da duplicidade de papéis do professor: não apenas transmissor mas também gerador de conhecimento.

A questão que se coloca então é como o professor de língua portuguesa pode gerar conhecimento, já que não há dúvida sobre seu papel de transmissor. Para isso, proponho aqui uma sugestão de roteiro.

### ROTEIRO PARA INICIAÇÃO EM PESQUISA

Um projeto de pesquisa em Língua Portuguesa pode partir de dois grandes aspectos: descrição ou uso da língua, apresentados aqui como dois extremos, intermediados por uma faixa de continuidade. A descrição envolve o extremo teórico da língua, tanto quanto possível desprovido de contexto; é a língua dentro dos limites da própria língua: da fonologia, do léxico e principalmente da



sintaxe. Questões de pragmática, levando em consideração, por exemplo, o objetivo para o qual se produz um determinado enunciado – ou questões de discurso, com preocupação no efeito de sentido de uma frase num determinado contexto sócio-histórico – tudo isso é ignorado. Geralmente não se vai além da frase, onde o significado fica restrito a uma forma lógica, envolvendo um falante ideal, desprovido de todas as circunstâncias sociais, ou mesmo de influências psicológicas, que possam afetar o enunciado.

No outro extremo está a língua em uso em situações autênticas de comunicação, levando em consideração tudo aquilo que acontece durante a fala, a escuta, a leitura e a escrita. A ênfase não está mais na frase, mas no evento comunicativo, que pode ser uma interação entre professor e alunos, uma consulta ao médico, uma entrevista para conseguir um emprego, um texto a ser lido, um relatório a ser escrito, etc. Leva-se em consideração não apenas a materialidade lingüística mas também os interlocutores, os papéis que eles representam, o lugar de onde eles produzem os enunciados, etc.

Quando se inicia um projeto de pesquisa já se está normalmente inserido em algum lugar desse contínuo que vai da descrição ao uso da língua. O importante aqui é dar-se conta de que minha maneira de ver a língua não é a única. Há também outros olhares, que não são nem superiores nem inferiores ao meu; são apenas diferentes. Daí que a atitude certa é não se deixar levar nem pelo desprezo das outras áreas, nem por um complexo de inferioridade.

Sabemos que a pesquisa é um procedimento sistemático de achar respostas para perguntas. Essas perguntas podem ser tanto teóricas como práticas, envolvendo a língua em todos os seus aspectos, desde a descrição mais pura até o uso mais contextualizado. Um exemplo de pergunta teórica: qual é o papel da sintaxe na compreensão da língua? Até que ponto a disposição das palavras na frase pode afetar sua compreensão? Qual é a diferença entre “O cachorro mordeu o menino” e “O menino mordeu o cachorro” (onde a única mudança é a ordem das palavras na frase)? Qual é a estrutura subjacente em “Pintou a parede de madeira” e em “Pintou a parede de vermelho”? Qual é a diferença entre “Posso estudar” e “Quero estudar”?

Um exemplo de pergunta prática, envolvendo não a descrição mas o uso da língua: Como escrevem os professores de língua portuguesa? Que tipo de revisão fazem quando reescrevem os próprios textos? Com que se preocupam mais? Com as questões locais da gramática? Com a coesão? Com a coerência do seu próprio texto?

Quando se procura uma resposta para uma pergunta, a primeira coisa a fazer é tentar descobrir se a pergunta já foi respondida; se já foi, talvez não haja razão para empreender o projeto. É preciso saber no entanto como a pergunta foi respondida, se baseada em pesquisa, em autoridade, ou mesmo em crenças pessoais. Mesmo que tenha sido baseada em pesquisa, é preciso saber qual foi a metodologia usada, isto é, quem foram os sujeitos que forneceram os dados, que procedimento foi usado para coletar esses dados, que tipo de análise foi feita, etc. Usando esses critérios, vamos logo descobrir que pouquíssimas são as perguntas que já foram respondidas. O problema quando se

inicia um projeto de pesquisa não é descobrir que outros já pesquisaram o mesmo assunto antes de nós; o problema mais comum é ver que praticamente nada foi feito e temos que começar da base, usando às vezes muito pouco conhecimento recebido.

Para saber até que ponto uma pergunta já foi respondida, fazemos o que se chama tradicionalmente de *fundamentação teórica*, mas que eu prefiro chamar de *revisão da bibliografia*. Aqui se entra numa tensão interessante: quanto menos bibliografia se encontrar para responder nossa pergunta, mais interessante e original deve ser nosso projeto, mas também mais difícil será desenvolvê-lo, porque não temos onde nos apoiar. Por outro lado, se encontrarmos muita bibliografia sobre o assunto, fica também mais difícil encontrar uma pergunta que ainda não tenha sido respondida, e assim criar um nicho próprio de pesquisa. A solução neste caso está em refinar a própria pergunta na medida em que se avança nas leituras; em qualquer área de conhecimento há sempre lacunas que precisam ser preenchidas.

Uma revisão bibliográfica hoje não pode ignorar os recursos da tecnologia da informação. Saber o que outros disseram sobre o assunto que queremos pesquisar é muito mais fácil agora do que há alguns anos. Há inúmeros bancos de dados, incluindo textos integrais de trabalhos de pesquisa, não só na Internet, mas também em forma de CD-ROMS. Muitos dos grandes periódicos acadêmicos, como o *TESOL Quarterly*, e *Theory into Practice*, têm também publicações eletrônicas de seus textos, reunindo em um único CD tudo o que foi publicado desde a primeira edição.

No Brasil, o CD-ROM TELA (Textos em Lingüística Aplicada), lançado em fins de 2000 (Leffa, 2000), que tive o prazer de compilar, reúne mais de 44.000 páginas de texto na área de Lingüística e Lingüística Aplicada. A figura 1 mostra o sumário do CD com sua classificação inicial. São oito tópicos de partida que se ramificam rapidamente em outras divisões, até levar o leitor, através de atalhos dinâmicos, ao texto procurado. São livros, periódicos da área, anais de congressos, teses de doutorado, dissertações de mestrado, além de inúmeros outros trabalhos, o equivalente a 220 livros de 200 páginas cada um, o que dá para encher uma estante.

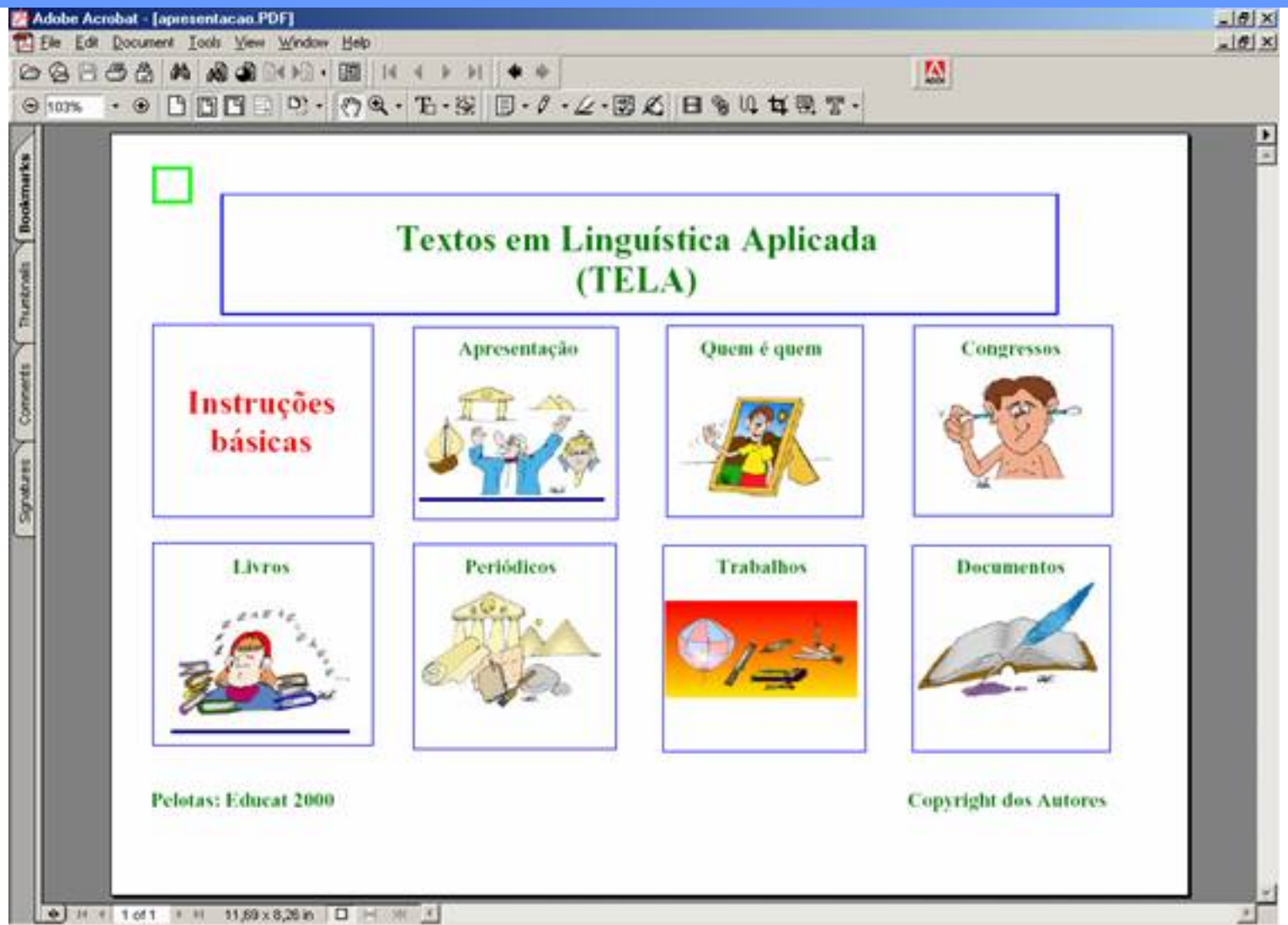


Figura 1 – Sumário do CD-ROM TELA (Textos em Linguística Aplicada)

Na medida em que as 44.000 páginas estão totalmente indexadas, é possível localizar em alguns segundos qualquer autor, qualquer tópico, e mesmo qualquer palavra de todo o acervo do CD. A figura 2 mostra dois momentos de uma busca específica, ou seja, procuram-se textos que contenham as palavras *ensino*, *pesquisa* e o segmento *portugu\** (seguido de asterisco). Isto significa que todos os arquivos de texto que contenham as palavras *ensino* e *pesquisa* e também todas as palavras que iniciam com o segmento *portugu* (português, portuguesa), estarão disponibilizados para o leitor – por ordem de proximidade entre as três palavras. O segundo momento mostra o resultado da busca, onde se vê que há 98 arquivos com os termos procurados, dinamicamente hierarquizados pelo critério da proximidade, de modo a mostrar em primeiro lugar os textos que provavelmente serão mais pertinentes para a pesquisa do leitor. O CD TELA é um bom lugar para se iniciar uma revisão bibliográfica sobre praticamente qualquer tópico da língua portuguesa, tanto teórico como prático.



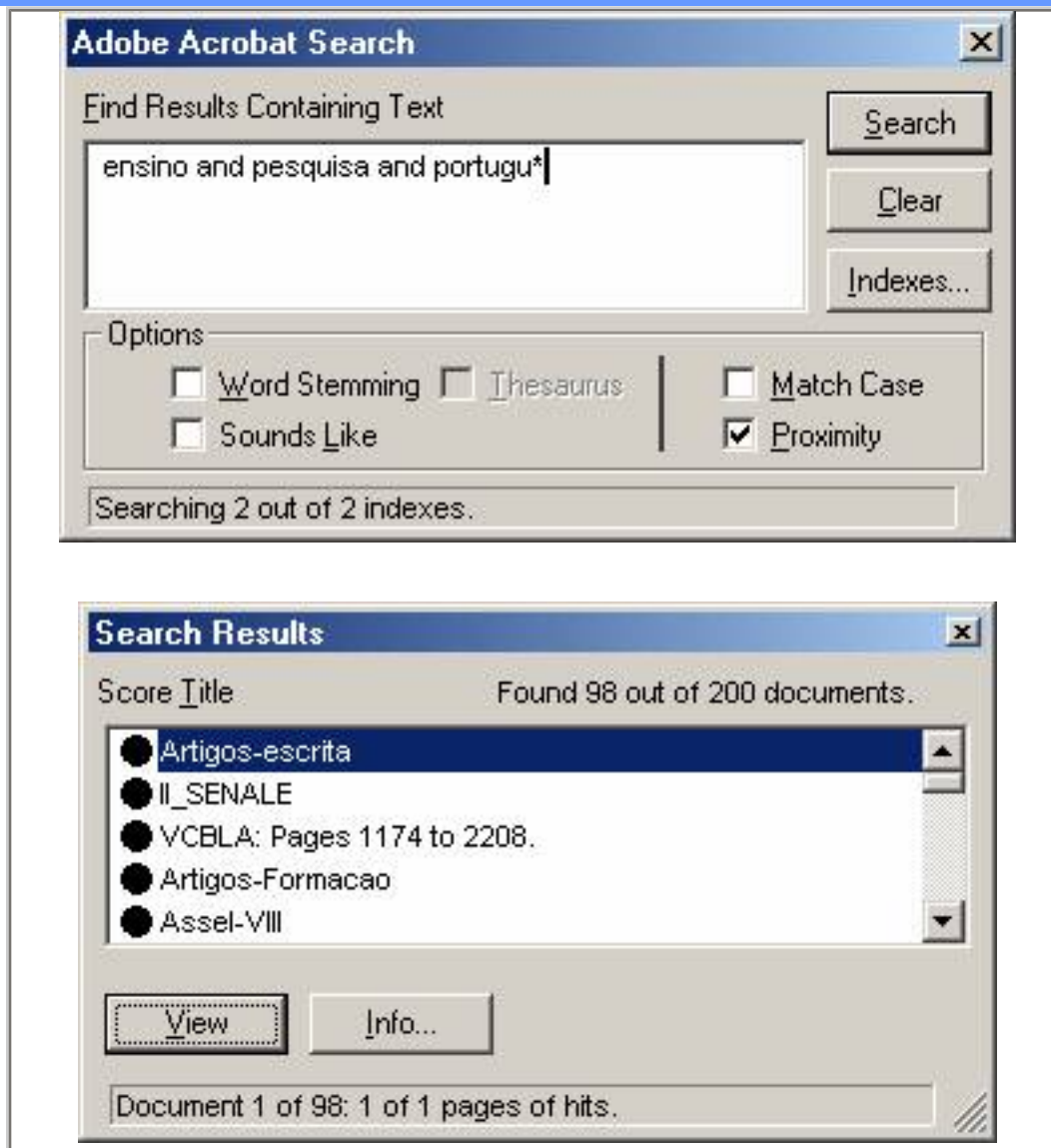


Figura 2 – Exemplo de uma busca feita no CD TELA.

Feita a revisão teórica, o passo seguinte é a coleta de dados. Também aqui temos recursos com os quais nem sonhávamos há algumas décadas. Além do recurso de gravações de áudio e vídeo, a preços economicamente viáveis (uma câmara de vídeo hoje custa uma fração do que custava há duas décadas), há também uma grande variedade de meios de comunicação de massa (TV a cabo, satélite, etc.), que podem fornecer inúmeros tipos de dados. O que mais chama atenção, no entanto, são os grandes bancos de textos disponíveis a qualquer pesquisador interessado no estudo da língua portuguesa, principalmente os corpora de textos de jornais e revistas (*Folha de São Paulo*, *Super Interessante*, *Exame*, etc.).

Figura 3 – Formulário para busca de ocorrências no CD Folha, Edição 99.

A idéia aqui não é usar esses textos como fonte de informação, mas como dados para pesquisa. Vamos supor, por exemplo, que alguém esteja interessado em saber como as pessoas que usam profissionalmente a língua portuguesa lidam com a regência do verbo *assistir*, no sentido de *comparecer*. De acordo com a gramática prescritiva o verbo *assistir* neste caso deve ser regido pela preposição *a*: *assistir ao* jogo e não *assistir o* jogo. Isso é o que diz a gramática, mas quem está interessado no estudo da língua em uso vai também estar preocupado com que as pessoas dizem. Todos esses bancos de texto, escritos por jornalistas, geralmente podem ser rastreados em termos de qualquer palavra ou mesmo segmento de palavra, como é o caso do CD da *Folha de São Paulo*. A figura 3, por exemplo, mostra que nas formas em que inicia por *assisti*, o verbo tem 12.576 ocorrências (edição 1999 do CD). A partir daí é só clicar o mouse e comparar as ocorrências, com e sem preposição. É muito provável que nem precise das 12.576 ocorrências, podendo adotar um critério para tirar uma amostra. A título de curiosidade, a Figura 4, mostra um exemplo em que o verbo não está regido por preposição (“O técnico Wanderley Luxemburgo fez poucas intervenções

da tribuna de rádio de onde assistiu o jogo”). (Também, a título de curiosidade, acho que seria interessante observar que o corretor do meu processador de texto assinalou que deveria escrever “assistiu ao jogo”, o que provavelmente vai contribuir para que essa forma se torne mais freqüente, dependendo da popularidade do processador).

Folha99

33 de 12576

ARTIGO

Título:	Técnico fala pouco no jogo
Data:	24/Dez/98
Autor:	Maércio Santamaria; Luiz César Pimentel; Fernanda Papa
Editoria:	Esporte
Página:	4-4
Edição:	Nacional
Seção:	
Assunto:	CAMPEONATO BRASILEIRO DE FUTEBOL, 1998; FINAL; CORINTHIANS /CLUBE/; CRUZEIRO /CLUBE/; WANDERLEY LUXEMBURGO
Observação:	SUB-RETRANCA

da Reportagem Local

O técnico Wanderley Luxemburgo fez poucas intervenções da tribuna de rádio de onde **assistiu** o jogo.

Antes, por meio de um pedido de liminar do Sindicato dos Treinadores Profissionais do Estado de São Paulo, ele tinha tentado assegurar meios para dirigir do campo no Corinthians, mas não conseguiu.

O treinador cumpre suspensão de 60 dias por ter sido expulso em jogo contra o Vitória, na primeira fase, após discutir com o árbitro Cláudio Cerdeira. No primeiro tempo, Luxemburgo falou, por um ponto eletrônico, com o auxiliar técnico Oswaldo Oliveira três vezes.

Pedia principalmente que Mirandinha e Edilson voltassem mais para marcar e apontou demora na armação dos contra-ataques.

No intervalo, no caminho do vestiário, não deu entrevistas. Na volta para o segundo tempo, depois de errar a cabine na qual acompanhava o jogo, disse que não era hora de falar.

Só se manifestou no segundo tempo, quando Edilson marcou o primeiro gol. A festa na cabine foi maior com o gol de Marcelinho. "A partir daí, não consegui mais ouvir o que ele dizia", disse Oliveira, que escutava Luxemburgo do campo.

texto integral

Figura 4 – Exemplo de ocorrência do verbo *assistir* sem preposição.

Os próprios dicionários eletrônicos podem também ser usados não só para pesquisar mas também para brincar, caso haja diferença entre *pesquisar* e *brincar*. Usando o recurso de “pesquisa reversa” do *Aurélio*, é possível descobrir inúmeros dados da língua portuguesa: entre tantos outros, o fato de que há no dicionário 25.273 brasileirismos e 2.420 palavras do tupi. Em termos de classes gramaticais, com alguns toques de mouse, descobrimos logo que há 82.406 substantivos, 40.555 adjetivos e 12.602 verbos – eu, pelo menos, até então não sabia que para cada verbo registrado no dicionário, há cerca de 7 substantivos. Como a pesquisa reversa permite que se parta da definição para a palavra definida, pode-se, com um pouco de tentativa e erro, encontrar a expressão de



qualquer conceito. Usando-se, por exemplo, as palavras *mulher*, *medo* e *aversão* pode-se chegar – a quem interessar possa – ao termos abaixo:

1. **Ginecofobia:** Aversão patológica a convívio com mulheres; ginofobia.
2. **Misoginia:** Desprezo ou aversão às mulheres.
3. **Partenofobia:** medo mórbido de mulher virgem

Um projeto de pesquisa começa com uma pergunta e continua com a exploração dos recursos disponíveis. A pergunta leva à hipótese e a necessidade de testar a hipótese leva à exploração dos recursos. O segredo do sucesso do pesquisador está em imitar as crianças, que, com sua curiosidade, aprendem a falar e, com alguma sorte, até a escrever. Sem curiosidade, sem o processo constante de tentativa e erro, não há pesquisa e nem avanço do conhecimento.

### A SEGMENTAÇÃO DA FRASE

Uma competência fundamental no processamento da língua é a capacidade de segmentar o que se ouve ou se lê. Quando estamos expostos ao fluxo da fala ou do texto escrito, temos que separar os períodos em orações, as orações em sintagmas, os sintagmas em palavras, as palavras em sílabas, as sílabas em fonemas ou letras, sendo talvez interessante observar que é só na última segmentação que temos um termo específico para a representação sonora (fonema) e outro para a representação gráfica (letra); sílabas, palavras, orações, por outro lado, existem tanto para a fala como para a escrita.

Ensinar o aluno a separar esses diferentes segmentos em outros menores tem sido sempre uma grande preocupação da escola, principalmente nos segmentos maiores como, por exemplo, na identificação das orações em um período composto.

Imaginemos a seguinte frase:

(1) José virá se amar Maria.

Temos aí um período composto de duas orações: uma oração principal, *José virá*, e uma oração subordinada adverbial condicional, *se amar Maria*. Qualquer falante da língua portuguesa será capaz de fazer essa segmentação, independentemente de conhecer ou não a terminologia adequada. A complicação, no entanto, pode surgir quando se invertem as orações no período, como em (2), propositadamente escrito sem vírgulas.

(2) Se amar Maria José virá.

Há duas complicações aqui. A primeira é causada pela ambigüidade do verbo *amar*, que pode ter mais de uma valência: valência 1, tendo como argumento apenas o sujeito, sem complemento: *José ama*. O verbo *amar* tem essa estranha propriedade de ser também um verbo intransitivo, onde não há necessidade de se especificar a quem se ama, como se amar fosse um sentimento sem direção e sem objeto.

Mas o verbo *amar* também pode ter valência 2, apresentando dois argumentos, sujeito e objeto: *José ama Maria*; ou mesmo sujeito e pronome: *José se ama*.

A outra complicação é que na língua portuguesa não há necessidade de explicitar o sujeito gramatical; posso dizer tanto *José virá* como simplesmente *virá*.

Isso significa que a frase *Se amar Maria José virá* pode ser segmentada de três maneiras diferentes:

(3) Se amar / Maria José virá.

(4) Se amar Maria / José virá.

(5) Se amar Maria José / virá.

Note-se que cada segmentação produz um sentido diferente. Em (3) quem ama é Maria José; já em (4) é José que ama Maria; enquanto que em (5) é alguém que ama Maria José. Não só é necessário segmentar para produzir sentido, mas é também necessário saber segmentar hierarquicamente, com flexibilidade, às vezes mais para a esquerda, às vezes mais para a direita, como se tivéssemos fronteiras móveis dentro do texto.

Trago esses exemplos para mostrar a importância da segmentação e da análise no processamento da língua. Pode-se afirmar com toda tranquilidade que não há compreensão e nem produção linguística se não houver processamento sintático. O estudo que vou resumir aqui (Leffa, 2000b) aborda essa questão da análise sintática de modo bastante simples e serve para mostrar que não é difícil implementar um projeto de pesquisa com resultados relevantes – apesar da simplicidade.

O objetivo do projeto foi tentar medir o papel da escola no desenvolvimento da capacidade de análise sintática do aluno. Para isso se preparou um teste que constou de 44 períodos compostos de duas orações, onde o aluno era solicitado, sem qualquer terminologia específica, a colocar um traço vertical onde achasse que a frase pudesse ser separada em duas partes. O teste foi aplicado em alunos da quarta série do ensino fundamental até alunos universitários do quarto semestre, passando pelo ensino médio. Os exemplos abaixo (6-9) mostram algumas das frases usadas.

(6) Acabada a festa, os músicos partiram

(7) Quero que vás mas peço que voltes



(8) Importa viver honestamente.

(9) A verdade é que todos o amam.

Os resultados mostraram três aspectos interessantes. Em primeiro lugar, notou-se que algumas orações eram mais fáceis de serem segmentadas do que outras: as orações separadas por vírgula (6) e as orações coordenadas (7) produziram mais de 90% de acertos; já as substantivas predicativas (9) foram as mais difíceis, produzindo menos de 20% de acertos.

Outro aspecto que me pareceu interessante foi a proximidade de resultados entre os extremos. O que era difícil para a quarta série do ensino fundamental era também difícil para o aluno universitário; e o que era fácil para o aluno universitário era também fácil para o aluno do ensino fundamental.

O terceiro aspecto, concluindo a pesquisa, é que não houve diferença significativa entre um extremo e outro: os alunos da quarta série do ensino fundamental obtiveram rigorosamente os mesmos escores que os alunos universitários; ambos acertaram, em média, 31 das 44 questões do teste. Isso significa que o ensino explícito da análise sintática, pressupondo, na verdade, que ele tenha ocorrido, não surtiu qualquer efeito na capacidade do aluno em segmentar as orações. Os resultados confirmam a hipótese de que a criança depois de uma certa idade não tem mais nada a aprender em termos de sintaxe – o que leva a conclusão de que o professor também não tem mais nada a ensinar. Na realidade, o que o professor não deve fazer é tentar ensinar ao aluno algo que ele já sabe e que ele, professor, não sabe explicitar.

Minha conclusão, em relação a este estudo, não é de que o professor de português não deva estudar gramática. Pelo contrário, deve estudar muito mais, incluindo as várias teorias que andam por aí (gramática de valências, Tesnière, gramática funcional, etc.), para então ser capaz de problematizar a língua para o aluno, dando-lhe a oportunidade de vislumbrar quão complexo é o conhecimento lingüístico que ele possui. Isso provavelmente é bem menos difícil do que parece, na medida em que se pode partir dos exemplos fornecidos pelos próprios alunos em seus textos e leituras, fazendo então a ponte com uma ou outra gramática.

### A CONEXÃO DAS PALAVRAS NA FRASE

O contrário de segmentar é conectar. Se por um lado separamos e decomposmos os elementos da frase, por outro, também unimos e conectamos. Os exemplos abaixo mostram como a conexão entre as palavras dá-se por estruturas subjacentes e podem mostrar a importância da sintaxe na compreensão:

(10) Pintou a parede de madeira.

(11) Pintou a parede de vermelho.

(12) O guarda viu a menina de binóculos.

(13) O guarda viu a menina de vestido vermelho.

No exemplo (10), podemos dizer que a palavra *madeira* está conectada à palavra *parede*, formando com ela um sintagma nominal, que, por sua vez, é o objeto direto do verbo *pintou*. Para gerar a frase, poderíamos fazer a pergunta: *O que ele pintou?* E teríamos a resposta: *Pintou a parede de madeira*.

Já em (11), a palavra *vermelho* se conecta não com a palavra *parede*, mas com o verbo *pintou*. Note-se que neste caso, para gerar a frase, perguntaríamos *Como ele pintou parede?*

A estrutura sintática da frase mostra também os casos de ambigüidade, em que mais de uma conexão é possível numa mesma frase. Em (12), por exemplo, a palavra *binóculos* pode estar conectada tanto à palavra *menina* como à palavra *guarda*, produzindo dois significados diferentes: (a) o guarda, que estava de binóculos, viu a menina; (b) o guarda viu a menina, que estava de binóculos.

Já em (13), a ambigüidade é mais difícil de ser percebida, por interferência do nosso conhecimento de mundo, onde normalmente não vemos guardas de vestidos vermelhos. A sintaxe, no entanto, mostra que a ambigüidade existe.

A sintaxe, é claro, não só mostra a ambigüidade onde ela potencialmente existe e às vezes nem percebemos, mas também resolve a ambigüidade onde a percebemos, como em (14), deixando imediatamente claro quando as palavras *ama* e *amo* são substantivos ou verbos.

(14) Amo a ama, mas a ama ama o amo.

A oposição entre dados estritamente lingüísticos – sintaxe e léxico – e a representação de mundo pode ser extremamente interessante nos estudos da compreensão, onde se procura verificar o peso de um e outro na determinação do significado. O conhecimento de mundo pode às vezes interferir no processamento lingüístico, como no exemplo da brincadeira, demonstrado em (15).

(15) O avião caiu na fronteira entre Canadá e Estados Unidos; onde devem ser enterrados os sobreviventes?

Dar à sintaxe o que é da sintaxe e ao conhecimento de mundo o que é do conhecimento de mundo tem sido um dos grandes debates nos estudos da compreensão. Se interpreto a leitura como um processo de extração de significado, estou provavelmente pondo ênfase na sintaxe, e indiretamente no léxico. Por outro lado, se interpreto a leitura como atribuição de significado, onde o leitor dá

sentido ao texto, estou enfatizando o que chamo de conhecimento de mundo, para não usar termos mais complicados.

Em termos de ligação, de retomada de partes do texto, um dos mecanismos mais interessantes na linguagem é a anáfora. A anáfora pode ser descrita como um processo que acarreta a volta no texto. O processo começa quando o anaforizante é reconhecido (por exemplo, o pronome) e concluído quando o anaforizado é encontrado (a palavra a qual o pronome se refere). Descrever o que acontece entre esses dois momentos, sem recorrer ao conhecimento de mundo, foi o que tentei no projeto que vou resumir em seguida. O objetivo foi verificar até que ponto as restrições sintáticas e lexicais do texto poderiam sozinhas contribuir para a resolução da anáfora.

O segmento abaixo, escolhido por sua simplicidade e pela ausência de ambigüidade, é usado para ilustrar os conceitos básicos subjacentes ao processo da anáfora, incluindo o rastreamento do antecedente.

(16) As **casas** são compradas porque **elas** oferecem conforto.

O pronome *elas* não se relaciona diretamente com um objeto no mundo real mas com palavras mencionada antes. A tarefa mental desempenhada pelo leitor ao processar a sentença, é voltar no texto e encontra a palavra a qual ele se refere. No Segmento (16), há cinco palavras (*porque compradas, são, casas e as*) mas somente uma é candidata séria para ser antecedente (*casas*). Vamos estabelecer, por enquanto, que o pronome *elas* só pode ser substituído por um substantivo plural feminino e a única palavra que preenche esta condição é *casas*.

É claro que na vida real as coisas nem sempre são tão simples. Um problema que pode surgir é a possibilidade de haver mais de um candidato legítimo para o antecedente, como está demonstrado no caso seguinte:

(17) As **casas** são compradas pelas **pessoas** porque **elas** oferecem conforto.

Agora não há um mas dois candidatos para *elas*, que são as palavras *casas* e *pessoas* (ambas substantivos plurais e femininas). Como resolver este problema? Uma hipótese é resolvê-lo aplicando as restrições sintáticas. Pode ser argumentado que há um paralelismo sintático entre o substantivo *casas* e o pronome *elas*, isto é : *casas* e *elas* estão na posição do sujeito em suas

próprias orações. A palavra *pessoas* por outro lado, embora substantivo plural não compartilha deste paralelismo com *elas*. Assim, entre os dois candidatos, escolhemos o substantivo *casas*.

Restrições sintáticas baseadas no paralelismo, no entanto, parecem funcionar bem apenas quando os exemplos são cuidadosamente escolhidos.

No Segmento 17, por exemplo uma simples mudança num item lexical pode reverter totalmente a relação entre o anaforizante e o anaforizado. Isso pode ser observado no próximo segmento onde *oferecem* foi substituído por *querem*.

(18) As **casas** são compradas pelas **pessoas** porque **elas** querem conforto.

Mais uma vez, existem dois candidatos exatamente como nos exemplos anteriores. Mas se aplicarmos a restrição sintática, como fizemos antes, escolhendo o sintagma nominal que está na posição do sujeito, chegaríamos a palavra *casas*, que obviamente é a escolha errada (*\*Casas querem conforto*). O paralelismo sintático, que tão eficazmente facilitou a escolha entre os 2 candidatos no segmento anterior, parece não funcionar mais.

O único candidato que pode ocupar legitimamente a posição ocupada por *elas* é *pessoas*. O outro candidato (*casas*) viola a restrição semântica: “*casas não querem coisas, somente pessoas querem coisas*”.

Assim, o paralelismo sintático é superado pelas restrições semânticas. Não basta o anaforizado possuir a mesma função sintática do anaforizante. Ambos devem também compartilhar o mesmo traço semântico.

As restrições sintáticas e semânticas, desse modo, não bastam para resolver os problemas associados à resolução da anáfora, como pode ser também observado nos seguintes casos:

(19) As **montadoras** venderam as **pickups** para as **firmas árabes** porque **elas** ofereceram cinco anos de garantia.

(20) As **montadoras** venderam as **pickups** para as **firmas árabes** porque **elas** eram à prova de bala.

(21) As **montadoras** venderam as **pickups** para as **firmas árabes** porque **elas** ofereceram pagamento à vista.

Os segmentos 19-21 aparentemente podem ser resolvidos apenas pela recorrência das representações de mundo em que compradores, vendedores e mercadorias trocam de mãos: dinheiro de compradores para vendedores e carros de vendedores para compradores. Precisamos saber também que carros podem ser a prova de bala, que as companhias oferecem garantias sobre o que elas vendem e que os árabes podem ser muito ricos.

Para este estudo sobre anáfora foram escolhidos os pronomes possessivos *seu, sua, seus, suas*. Esses pronomes apresentam algumas características que os tornam interessantes para uma investigação sobre a resolução da anáfora.

A mais importante é que no Português esses pronomes não concordam com o possuidor, como acontece no Inglês, mas com a coisa possuída. A principal consequência desta concordância pelo lado direito (com a coisa possuída) é que o rastreamento do antecedente se torna mais difícil. Não se podem usar, por exemplo, pistas sintáticas importantes, como gênero e número.

Pressupõe-se que todas essas dificuldades sobrecarregam a hipótese de que apenas as restrições sintáticas e semânticas, sem conhecimento de mundo, possam levar a uma resolução da anáfora.

Resultados positivos, obtidos em condições tão desfavoráveis, serão provavelmente mais robustos e confiáveis.

A metodologia usada para testar a hipótese consistiu de um levantamento de um corpus de 1.300 ocorrências de pronomes possessivos em textos jornalísticos. A base para a obtenção desses dados foi o CD-ROM da *Folha de São Paulo*.

Quadro 1 – Restrições na resolução da anáfora

Tipo de restrição	Percentual
Restrições sintáticas (sujeito)	85%
Restrições semânticas	10%
Não resolvido	5%

O quadro 1 mostra os resultados obtidos. Em 85% dos casos o anaforizado é o sujeito da oração principal. Em (22), por exemplo, teoricamente há três candidatos para o antecedente de *seu* em *seu pai*, mas apenas *Romário* preenche a condição de sujeito da oração.

(22) **Romário** disse ontem que **traficantes** ajudaram a encontrar o **cativoiro** de **seu** pai, libertado



anteontem.

Incluindo as restrições semânticas e considerando as palavras que enfatizam o objeto como os verbos *tratar de*, *homenagear*, *quanto ao*, *segundo*, etc. é possível aumentar o índice de resolução em mais 10%. Os exemplos abaixo mostram como a anáfora funciona nesses casos (Note-se que em nenhum dos casos o antecedente é o sujeito da oração principal):

Quanto aos **vira-latas**, **Coren** diz que a **aparência** é o que mais vale na **hora** de adivinhar **seu** temperamento

O **livro** trata do **indivíduo** e da **necessidade** de encontrar **seu** destino.

Segundo **Faria**, um **carro azul**, **quatro portas**, parou a **seu** lado, com três homens. Um deles saltou, apontando-lhe uma arma. Faria foi atirado dentro do carro.

**Maluf** disse que vai homenagear **Senna** colocando **seu** nome em alguma obra pública.

A conclusão é de que não existe compreensão sem processamento sintático, tanto em termos de segmentação, onde é preciso identificar os elementos que compõem a frase e o texto, como em termos de conexão, onde é preciso ligar, ou religar, o que foi separado para análise. Se eu afirmar que a compreensão envolve dois movimentos opostos de análise e síntese, não estou dizendo nada de novo, mas depois dos dois estudos realizados e das reflexões que eles proporcionaram, sinto-me mais próximo do problema. Tenho mais tranquilidade em afirmar que o conhecimento de mundo é importante, mas que o processamento sintático não pode ser ignorado na compreensão do texto.

## RECOMENDAÇÕES FINAIS

O objetivo desta apresentação foi argumentar que o professor não deve apenas passar conhecimento para o aluno, mas também produzir conhecimento. Apresentei aqui, a título de ilustração, dois exemplos de pesquisa que podem ser desenvolvidos com o que temos na sala de aula ou em bases de dados disponíveis aos professores curiosos ou interessados. Um estudo foi mais prático, envolvendo um questionamento do ensino da análise sintática; outro, mais teórico, envolvendo a anáfora. Nenhum dos dois vai salvar o mundo, mas pesquisa é assim mesmo. O importante é que cada um faça sua parte.

Algumas recomendações finais:

Tente produzir pelo menos parte do conhecimento que você passa para o aluno. Pesquise como seus alunos reescrevem seus textos, que estratégias você poderia usar em sua sala de aula para dar um leitor para seu aluno, que tipos de textos eles gostariam de ler. Essa são apenas alguns exemplos de tópicos de pesquisa que precisam ser investigados.

Tente descobrir outros colegas que tenham interesses semelhantes aos seus. Hoje é possível formar comunidades discursivas sem limites geográficos, para a troca de idéias e experiências. Um professor sozinho não faz pesquisa; é preciso formar uma comunidade.

Tente refletir sobre a experiência de ensino. A sala de aula é um verdadeiro laboratório onde muitas coisas podem ser experimentadas e o professor é de certa maneira um privilegiado por dispor de todos esses dados. Mas não basta viver toda essa experiência; é preciso refletir sobre ela.

Vou concluir repetindo aqui o que já disse em outro lugar:

O conhecimento é a matéria prima do professor, que deve agir não apenas como transmissor de saber, mas também como produtor. É difícil saber o que é mais importante, se gerar ou transmitir conhecimento. A preocupação em estabelecer essa diferença, no entanto, não existe para o professor, na medida em que ele pode exercer as duas atividades. Se o conhecimento, por si só, nunca foi tão valorizado como agora, o professor, no duplo papel de gerador e transmissor de conhecimento, deve ser duplamente valorizado. Em muitos momentos da história, e em muitos setores, ainda hoje, os detentores do saber, muitas vezes, têm-se caracterizado por reter e até sonegar o conhecimento. O professor vai mais longe: distribui o conhecimento que possui. A importância maior do professor não está em construir o conhecimento para si, e nem mesmo em construir para os outros; a importância do professor está em construir o conhecimento nos outros (Leffa, 2001, p. 18).

## REFERÊNCIAS BIBLIOGRÁFICAS

FERREIRA, Aurélio Buarque de Holanda. *Novo Aurélio século XXI*; O dicionário da língua portuguesa / Aurélio Buarque de Holanda Ferreira. 3.ed. totalmente revisada e ampliada. Rio de Janeiro: Nova Fronteira, 1999.

FERREIRA, Aurélio Buarque de Holanda. *Novo dicionário da língua portuguesa*. 2.ed. revista e aumentada. Rio de Janeiro: Nova Fronteira, 1986.

FERREIRO, Emilia. *Cultura escrita e educação*; Conversas de Emilia Ferreiro com José Antonio Castorina, Daniel Goldin e Rosa María Torres. Porto Alegre: ARTMED, 2001.

LEFFA, Vilson J. (Compilador). *TELA*; Textos em Lingüística Aplicada [CD-ROM]. Pelotas: Educat, 2000a.

LEFFA, Vilson J. O ensino das linguas estrangeiras nas comunidades virtuais. Palestra apresentada no *IV Seminário de Línguas Estrangeiras*. Goiânia: UFG, setembro de 2001.

LEFFA, Vilson J.. Análise sintática; ensinando o que não se sabe para quem já sabe. In: FORTKAMP, Mailce B. M.; TOMITCH, Leda, Maria Braga. *Aspectos da Lingüística*

*Aplicada*; estudos em homenagem ao Professor Hilário Inácio Bohn. Florianópolis: Insular, 2000b. p.203-228.

MCKEON, Denise. Future perfect: Lessons for a 21st Century TESOL. *TESOL Matters*, USA, v. 11, n. 3, p. 1,4, 2001.