

O teor convencional do fenômeno lingüístico: quando o corpus fala mais alto

Milena Uzeda-Garrão*

RESUMO

Este artigo apresenta a importância do olhar convencionalista sobre o fenômeno lingüístico. Traz algumas breves considerações filosóficas sobre essa linha de pensamento; apresenta o seu desdobramento dentro da Lingüística e mais especificamente no domínio de Processamento de Linguagem Natural. Mais adiante, contrasta esse olhar àqueles mais notoriamente empregados para lidar com o fenômeno lingüístico, quais sejam, o universalismo chomskyano e o cognitivismo lakoffiano. Finalmente, discute a viabilidade e a acuidade de uma visão eminentemente convencionalista para lidar com o fenômeno lingüístico.

Palavras-chave: Teorias Lingüísticas, Convencionalismo, PLN, Lingüística de *Corpus*

ABSTRACT

This paper presents the importance of dealing with linguistic phenomenon from a conventionalist perspective. It makes a brief philosophical account on this theoretical framework; then, it presents its development within Linguistics and, more precisely, within Natural Language Processing. Finally, it contrasts this framework with those more traditionally applied in Linguistics, namely, Chomsky's Universalistic accounts and Lakoff's Cognitivist ideas. Finally, it discusses the feasibility and reliability of the conventionalist framework on linguistic phenomenon.

Keywords: Linguistic Theories, Conventionalism, NLP, Corpus Linguistics

Introdução

*O discurso é um grande soberano, que com o
mais diminuto e inaparente corpo as mais divinas
obras executa. (Górgias, Elogio a Helena)*

Neste artigo me aventuro por uma vereda relativamente pouco trilhada nos estudos lingüísticos, mas que já havia sido sinalizada pelas considerações sofistas da Grécia antiga. Trata-se de um olhar sobre o fenômeno lingüístico que privilegia o seu estatuto convencional. Dispensa, em contrapartida, qualquer tipo de teor simbólico ou representativo da linguagem em relação à realidade ou à mente.

Reservo a primeira seção do artigo para breves considerações filosóficas sobre essa linha de pensamento; apresento na segunda seção o seu encaminhamento no âmbito lingüístico e no âmbito de Processamento de Linguagem Natural (PLN,

* Instituto Militar de Engenharia Departamento de Engenharia de Computação

** PUC-Rio, Departamento de Letras.

doravante); já na terceira seção, contraste esse olhar àqueles mais notoriamente empregados para lidar com o fenômeno lingüístico, apresentando a vantagem ou, talvez seja melhor dizer, a conveniência desse ponto de vista se comparado a essas perspectivas mais hegemônicas no domínio teórico dos Estudos da Linguagem.

1. Da herança filosófica

Não é tarefa simples caracterizar as considerações convencionalistas sobre a linguagem. Na história do pensamento filosófico, a visão sofista sobre o discurso é talvez a primeira manifestação documentada dessa concepção de linguagem. Muito do que sabemos sobre o pensamento sofista em relação à linguagem deve-se aos escritos platônicos.⁶ O que, de certa forma, é sintomático, uma vez que Platão se opunha à visão convencionalista.

Os registros de Platão apresentam as considerações sofistas — em que a linguagem é vista à luz da célebre máxima de Protágoras, “o homem é a medida de todas as coisas” — como contraponto ao pensamento socrático. Este formato dialógico apresenta, de certa forma, um teor didático para reafirmação de sua crença representacionista. Em outras palavras, o fato de Platão sustentar, em oposição violenta ao ideário sofista, a existência de verdades únicas e fixas, em alguma medida pode ter minimizado ou desqualificado a importância e teor das considerações sofistas ao longo da história do pensamento filosófico (cf. Souza Filho, 1997 e Martins, 2003). Logo, esse investimento de Platão contra os sofistas deve ter contribuído para minimizar e desqualificar também o ângulo estatístico sobre a linguagem.

Dentre as características privilegiadas por essa visão de linguagem está a mutabilidade do significado de uma expressão em virtude de esta depender das práticas humanas, como também a incapacidade de a expressão representar algo exterior à linguagem.

Esse viés eminentemente pragmático sobre a linguagem é retomado já na história mais recente da filosofia⁷, notadamente, por Ludwig Wittgenstein⁸. De forma ostensivamente resumida, pode-se dizer que o filósofo austríaco se recusa a enxergar o estudo da linguagem dentro dos mesmos moldes metafísicos propostos para as ciências, principalmente porque o “cientista”, neste caso, é uma das peças envolvidas no que chama de “jogos de linguagem” (Wittgenstein, 1979; Glock, 1996). Ele desenvolve este conceito através de uma analogia com a idéia de “jogo”, e aponta certas características e semelhanças com a linguagem: jogos possuem regras, são práticas compartilhadas por uma comunidade, possuem peças, são autônomos, não requerem justificativas.

Sua proposta é a de que os jogos de linguagem “são a totalidade da linguagem e das atividades com as quais está interligada” (Wittgenstein, 1979: §§7 e 23). Os jogos de linguagem seriam atividades autônomas que prescindem de explicação; “são parte de nossa história natural, assim como andar, comer, jogar, etc”. (ibidem: §25). E o fato de o próprio homem estar encarcerado nos rituais lingüísticos que regem toda e qualquer manifestação lingüística compromete o seu

⁶ O *Sofista* e *Crátilo* são exemplos de diálogos em que Platão tematiza a linguagem.

⁷ Para uma boa apreciação sobre o assunto ver Martins, 1999.

⁸ Refiro-me ao segundo Wittgenstein e sua visão pragmática expressa em *Investigações Filosóficas*, São Paulo Cultural, Coleção *Os Pensadores*, 1979.

juízo teórico ou sua tentativa de explicação acerca do assunto.

O estudo sobre a linguagem, portanto, seria ele mesmo mais um desses jogos. E como o significado de uma palavra encontra-se na execução da linguagem, não há nada a ser provado ou justificado. Não se pode explicar o que está explícito. A significação é habilidade em lidar com as palavras na linguagem. A explicação é resultado da nossa "ânsia de generalidade". Ela pode até ser utilizada, mas não deveria ser entendida como uma meta-regra; ou uma regra fora do jogo. Toda explicação é interessada. (ver Scherer, 2002 sobre o conceito de *explicação*).

Sua visão nega, deste modo, a vocação representacionista da linguagem tão defendida por filósofos como Platão, Aristóteles e Locke. Não há, para ele, uma essência do significado que preceda o uso das palavras. É somente na prática do uso que o significado se dá. E mesmo assim, ele não poderia ser "coisificado". Por essa razão, é instável, contingente e fragmentado. Falar sobre a linguagem, ou fazer um relato sobre o relato, será um relato do jogo, e também faz parte dele. Portanto, nunca será definitivo.

Se num primeiro momento sua visão parece amputar a Lingüística, e mais especificamente, eliminar a Semântica, uma vez que sugere uma total impossibilidade de tratamento do fenômeno da significação, concordo com Martins (1999:147) quando pondera que adotar uma visão wittgensteiniana na Lingüística "não corresponde à negação da possibilidade de qualquer estudo sistemático e empírico das línguas do mundo". A autora sugere, portanto, que uma Lingüística sob a perspectiva wittgensteiniana "é viável e teria como propósito geral descrever as regularidades — parciais e contingentes — observáveis nos jogos de linguagem que constituem as línguas do mundo" (Idem). A descrição de uma língua, segundo o filósofo, estará sempre condicionada à parcialidade imposta pelas infundáveis possibilidades de lances dos jogos de linguagem.

2. Ecos da visão convencionalista na Lingüística e em PLN

Na falta de uma explicação satisfatória para a noção do significado, os lingüistas que atuam na área da Semântica encontram-se na situação de não saber do que é que estão falando (W. Quine)

A posição do lingüista Roy Harris (1981), em relação ao funcionamento da língua alinha-se às idéias convencionalistas solidificadas por Wittgenstein. Em Scherer (2002) é possível encontrar uma farta exposição da simpatia de Harris às idéias do filósofo austríaco. Um dos pontos em comum entre as duas posições é a impossibilidade de distinção entre semântica e pragmática, ou entre conhecimento lingüístico e enciclopédico.

Note-se que a visão pragmática oferecida aqui se distancia de muitas abordagens pragmáticas da Lingüística. Estas conduzem suas explicações sobre o "uso da língua" com base na distinção Semântica-Pragmática, que com freqüência se apóia na distinção significado literal-significado de uso; algo indissociável por uma visão wittgensteiniana. Scherer (2002:26) avalia, portanto, que Wittgenstein e Harris compartilham uma visão pragmática radical, em que o uso lingüístico não é um

dos componentes da linguagem, mas a única forma produtiva de se pensar os fenômenos lingüísticos.

Adam Kilgarriff, estudioso do léxico do ponto de vista computacional, ou seja, com interesse voltado para PLN, também advoga este pragmatismo radical. Em seu contundente artigo "I don't believe in word senses" (2000), ele atribui ao corpus o poder de desambiguação de significados. Demonstra que as palavras, em suas diversas acepções, estão desatreladas do seu alegado "sentido atômico" (core meaning).

Propõe, como alternativa, uma ontologia, ou um conjunto de palavras semanticamente relacionadas, com base em uma convergência (cluster) estatística dos seus vários usos em corpus, em detrimento do seu sentido atômico. Os extratos do corpus são agrupados em sentidos específicos de acordo com os objetivos de tarefas específicas. Acredita, portanto, que os significados só existem dependentes de propostas ou tarefas (tasks). Ele conclui que sem um fim ou uma tarefa, os sentidos das palavras não existem.

2.1 Implicações de uma abordagem com base em *corpus*

Como o leitor já pode supor a esta altura do artigo, a descrição de uma língua, segundo Wittgenstein, estará sempre condicionada à parcialidade imposta pelas infundáveis possibilidades de lances dos jogos de linguagem. A propósito, sobre a idéia de completude da linguagem, Wittgenstein confronta o leitor com o seguinte questionamento no §18 de *Investigações Filosóficas*:

"[...] (e com quantas casas ou ruas, uma cidade começa a ser cidade?) Nossa linguagem pode ser considerada como uma velha cidade: uma rede de ruas e praças, casas novas e velhas, e casas construídas em diferentes épocas; e isto tudo cercado por uma quantidade de novos subúrbios com ruas retas e regulares e com casas uniformes".

Esse parágrafo de *Investigações Filosóficas* é particularmente relevante para legitimar uma escolha teórica e também metodológica comumente utilizada na Lingüística: uma abordagem com base em *corpus*. Um exemplário da língua sempre pode ser maior, é verdade; mas nunca deixará de ser um extrato válido do seu uso. Portanto, a idéia de abarcar a língua de uma forma totalizante seria uma utopia. Conclui-se, daí, que a parcialidade é inevitável.

Ao longo do tempo, algumas expressões são construídas, outras demolidas. Isto é parte do jogo. Todo *corpus* de uma língua seria, portanto, uma caracterização válida da mesma⁹.

Vejamos o caso do estudo lexicográfico para o francês de Verlinde & Selva (2001), que compararam a abordagem tradicional da construção de um dicionário de aprendizes de segunda língua, baseada na intuição do lexicógrafo, a uma abordagem baseada em *corpus*.

De um modo geral, verificaram que é atribuído ao número de entradas (macroestrutura) de um dicionário uma importância maior do que ao conteúdo de

⁹ para uma melhor apreciação das idéias wittgensteinianas ver Glock, 1996; Martins, 1999.

cada entrada (microestrutura). Portanto, a ênfase recai nas palavras simples em detrimento das combinações de palavras, o que vem a ser um contra-senso, já que, para fins comunicativos, os aprendizes necessitam muito mais de informações sobre combinações do que sobre palavras isoladas.

Os pesquisadores puderam detectar, por exemplo, que, mesmo com a difundida repressão do governo francês aos anglicismos, o *corpus* apresentava uma frequência bastante representativa desses estrangeirismos. A partir dessa constatação, podemos atribuir aos dados lingüísticos um poder maior do que geralmente lhes é conferido.

Por outro lado, os autores utilizaram como fontes periódicos jornalísticos, o que, segundo Sinclair (1991), pode ser questionável: "um *corpus* é uma coletânea de um material amplamente homogêneo, mas retirado de fontes diversas de forma que a individualidade de uma fonte se perde, a não ser que o pesquisador queira isolar um texto em particular (p.17). A diversidade de fontes é uma garantia de segurança dos dados" (p.18).

Corroborando as preocupações de Sinclair (1991), Verlinde & Selva (2001) destacam que a lexicografia de *corpus* é uma evidência empírica necessária à intuição do pesquisador, que serviria para preencher as lacunas de *corpora* que chamam de "não equilibrados". Portanto, eles ressaltam a urgência da construção de um *corpus* equilibrado para o francês e outras línguas.

Pode-se notar, portanto, que grande parte dos lexicógrafos que se dedicam ao estudo de *corpus* o sentimento de que a compilação de fontes diversas evita o paradoxo do *corpus*. Isto é, evita a formação de um *corpus* tendencioso. Contudo, essa caracterização do *corpus* ideal ignora a imagem da "cidade" apresentada por Wittgenstein e revela uma ilusão corrente na Lingüística: a idéia de que é possível fugir do paradoxo do *corpus*.

A questão freqüentemente intocada é a de que a língua em si não é completa; sempre é possível acrescentar mais uma casa ou mesmo uma rua, o que torna um *corpus* um fragmento de algo já fragmentado. Portanto, não há como fugir do paradoxo do *corpus* uma vez que a completude da língua também é algo inatingível.

Tal constatação, entretanto, não desabona um estudo descritivo da língua; o fato de não ser possível exaurir todas as possibilidades de jogos da linguagem não impede um lingüista ou um lexicógrafo de descrever o que geralmente é constatado nos jogos (cf. Martins, 1999:147). Por outro lado, é desejável que o pesquisador se cerque de alguns critérios para aferir a adequação de um *corpus* em função dos objetivos pragmáticos que se deseja alcançar.

2.2 O teor estatístico do fenômeno lingüístico

Embora proposta antes do advento de *corpora* computadorizados, a Teoria Contextual do Significado proposta pelo lingüista britânico J. R. Firth (1957), subjaz a uma abordagem a partir de *corpus*. Ele propõe que o estudo do significado e do contexto devem ser centrais para a lingüística. Recusa-se a admitir qualquer tipo de distinção entre *langue* e *parole* (estabelecida antes dele por Saussure) ou entre

competência e desempenho (estabelecida mais tarde por Chomsky), já que não enxerga a língua como uma entidade autônoma. Os eventos lingüísticos, nos seus termos, são recorrentes e repetidamente observáveis.

No mesmo ano em que Firth profetizou “*You shall know a word by the company it keeps*”, Chomsky tornou pública, em *Syntactic Structures*, a sua crença de que a língua deveria ser analisada em um nível muito mais profundo do que sua estatística superficial. Isto é, numa crítica à visão empirista, argumentou que uma abordagem com base em dados da língua consideraria como identicamente remotas as sentenças “*Colorless green ideas sleep furiously*” e “*Furiously sleep ideas green colorless*”. Ou seja, em qualquer modelo estatístico, as duas sentenças seriam igualmente excluídas, embora, como falantes, nós saibamos que a primeira obedeceria a algumas regras gramaticais. O que Chomsky defendia é que a primeira não poderia ser totalmente desconsiderada e equiparada à segunda simplesmente pelo fato de sua probabilidade de ocorrência ser remota¹⁰. Essa crítica, conhecida como *problema dos dados esparsos*, foi altamente influente para uma mudança de perspectiva teórica nos anos 60.

Mas a evolução de técnicas estatísticas aliada à possibilidade de construção de *corpora* mais robustos de alguma forma equacionou aquilo que Chomsky utilizava como emblema de sua crítica. Além disso, a forma de Firth enxergar a língua não se presta à comparação com o modelo chomskyano. Para ele a língua é um evento, uma forma de ação (*a way of “doing things”*); e é por isso que seria um ato legítimo por parte do lingüista se ater aos eventos discursivos propriamente ditos. Firth acreditava que o evento era um fim em si mesmo e não uma forma de acessar o “verdadeiro” objeto de estudo: o sistema subjacente a ele, como Chomsky advogava.

3. A confiabilidade de uma abordagem estatística em relação a outras perspectivas

Nesta altura, parece-me prudente esboçar dois outros olhares teóricos talvez mais correntes na Lingüística: a visão universalista e a visão cognitivista sobre o fenômeno lingüístico. Sobre a primeira, pode-se dizer com certa segurança que foi altamente influenciada por pensadores como Platão, Aristóteles e Locke.

Com efeito, a oposição entre significado literal e figurativo se apresenta em virtude de uma compreensão do significado como entidade. A certeza de que é possível através da linguagem estabelecer relações de verdade sobre as coisas impulsiona esses pensadores a se aterem ao campo da literalidade. Um discurso figurativo, portanto, seria identificado com falta de clareza e imprecisão; careceria do poder epistemológico presente somente no discurso literal.

Portanto, o papel acessório que, tradicionalmente, vem sendo reservado à figuratividade na lingüística tem uma hegemonia histórica no pensamento filosófico. Afinal, é de Aristóteles a frase “*a metáfora consiste em dar à coisa um nome que pertence a outra coisa*”. E uma vez que a ciência e todo o pensamento ocidental são, em grande parte, herdeiros das suas considerações filosóficas, não é de se

¹⁰ Ironicamente, a probabilidade de ambas as frases ocorrerem é grande em virtude de sua relevância ilustrativa na Lingüística.

estranhar que o mesmo seja dito por muitos cientistas contemporâneos da área da linguagem, como, por exemplo Noam Chomsky e Steven Pinker.

Chomsky assim como seus legatários edificaram suas teorias neste pressuposto universalista. Toda manifestação lingüística não-literal, ou seja, não-representativa, deixa de ter importância do ponto de vista científico. De uma forma ampla, podemos dizer que, por esse olhar:

- a) O significado é uma propriedade exclusiva das expressões lingüísticas e se define em termos de referência e verdade assim como independem dos indivíduos que o produzem;
- b) É possível estabelecer o potencial referencial da linguagem; ou seja, como os símbolos se relacionam com a realidade;
- c) Deve-se ater exclusivamente ao domínio literal, não enciclopédico, não pragmático, declarativo e sincrônico;
- d) A linguagem tem prioritariamente a função de expressar (racionalmente) o mundo e, portanto, é possível estabelecer o valor de verdade das sentenças.

Já a visão **Cognitivista** sofreu forte influência de pensadores mais relativistas, como, por exemplo, Friedrich Nietzsche, um dos críticos mais afiados à filosofia tradicional.

Para Nietzsche, a língua é um grande depósito de metáforas e metonímias e o impulso à verdade se origina da necessidade de evitar "a guerra de todos contra todos": um "acordo de paz"(1978:46) do qual a humanidade não consegue se desvencilhar: "Não possuímos nada mais do que metáforas das coisas, que de nenhum modo correspondem às entidades de origem".

Volto-me agora à sua ressonância na visão lingüística proposta por George Lakoff. Segundo a sua proposta, os sistemas conceituais variam, ou seja, não haveria uma visão de mundo unicamente verdadeira; a gramática não seria pura forma; a emoção teria conteúdo conceitual; a mente não seria uma dimensão independente do corpo, e o significado não seria baseado em verdade e referência ou na relação entre os símbolos e as coisas no mundo.

Lakoff introduz o paradigma do **experencialismo** baseando-se tanto nos aspectos imaginativos quanto nos aspectos corporais da cognição. Argumenta que a visão essencialista da linguagem está equivocada ao afirmar que as categorias são universais e que existe uma linguagem do pensamento capaz de captar essas categorias uniformemente (*o mentais*, cf. Pinker, 1995:cap. 3). Para ele, as categorias lingüísticas são claramente reflexos da experiência, da imaginação e do corpo.

De uma forma ampla, mas talvez bastante elucidativa, podemos traçar características definidoras dessa linha de pensamento:

- a) O significado não é puramente lingüístico; pertence ao escopo das estruturas cognitivas gerais fundadas na nossa experiência concreta

(universal/cultural); a linguagem é simplesmente uma de suas manifestações superficiais.

b) Cabe ao lingüista descobrir e explicar a estrutura conceptual que subjaz à linguagem;

c) O escopo de análise é, portanto, muito mais amplo do que aquele proposto pelo representacionismo: inclui-se o não literal, o enciclopédico, o pragmático e o diacrônico;

d) Não há uma ambição verificacionista; não é uma semântica calcada em lógica abstrata;

e) A característica principal da linguagem é servir de sinal concreto para "complexas construções não lingüísticas"

f) Há ênfase nos processos humanos de categorização, na polissemia, e no poder figurativo da linguagem.

Por uma perspectiva puramente **convencionalista**, contudo, algumas dessas grandes questões epistemológicas não são simplesmente deixadas de lado. Apenas não se prestam a uma verificação incontestável como em outros campos do saber. Tal escolha leva a uma perspectiva lingüística amplamente amparada pela faceta estatística do significado. Trata-se de uma forma de enxergar as recorrências lingüísticas abrindo mão de rotulações semânticas pré-concebidas ou de considerações experientialistas.

Essa visão nega, deste modo, a vocação representacionista da linguagem tão defendida por filósofos como Platão, Aristóteles e Locke. Não há, para ele, uma essência do significado que preceda o uso das palavras. É somente na prática do uso que o significado se dá. E mesmo assim, ele não poderia ser "coisificado". Por essa razão, é instável, contingente e fragmentado. Falar sobre a linguagem, ou fazer um relato sobre o relato, será um relato do jogo, e também faz parte dele. Portanto, nunca será definitivo.

Portanto, Firth, Kilgariff e Harris fogem de certa forma desse embate aparentemente infundável entre universalismo e relativismo. Tal escolha de um caminho minimamente comprometido com representação de significado, contudo, não se trata de uma fuga ou uma evitação do confronto entre universalismo e relativismo, apenas de uma clareza de que tal confronto talvez não chegará a uma resposta definitiva.

De fato, não parece nenhuma heresia afirmar que a língua pode ser descrita como um fenômeno probabilístico, uma vez que há nitidamente padrões de combinações vocabulares recorrentes. Esses padrões, muitas vezes, extrapolam o nível lexical (como expressões idiomáticas mais triviais), culminando em estruturas sintagmáticas super-oracionais (Ver Jackendoff, 1997, cap. 7 e Uzeda-Garrão & Dias, 2001, para exemplos em inglês americano e português do Brasil, respectivamente).

De certa forma, esta perspectiva atenua a visão chomskiana da linguagem, focada no poder criativo do falante ideal, e prioriza uma visão de língua inseparável da pragmática; isto é, enfatiza o teor eventivo do fenômeno lingüístico.

Hoje em dia já dispomos de bons recursos estatísticos para detecção de

estruturas convencionais da língua. Existem alguns métodos estatísticos disponíveis para a detecção de convenções lingüísticas. Por isso, o avanço tecnológico teve um papel seminal na constatação de que uma visão probabilística da língua é viável e altamente profícua. E o *corpus*, além de servir como base de dados para identificação de tais convenções, também tem um papel preditivo ao fornecer os ambientes lingüísticos tipicamente relacionados a elas.

Como Wittgenstein define em *Da certeza* (§559), "o jogo de linguagem é, por assim dizer, imprevisível. Quero dizer: não está fundamentado. Não é racional (ou irracional). Está aí - como a nossa vida". Portanto, como jogadores, talvez a atitude mais prudente seja a constatação e descrição de partes dos jogos, sem tentar alçar vôos teóricos mais ambiciosos.

Referências

GARRÃO & DIAS. (2001/2) "Um estudo de expressões cristalizadas e sua inclusão em um tradutor automático bilíngüe". **Cadernos de Tradução no. VIII**, UFSC.

GLOCK, H.J. (1996) **Dicionário Wittgenstein**. Rio de Janeiro: J. Zahar.

HARRIS, R. (1996) **The Language Connection**. Thoemmes Press. UK.

JACKENDOFF, R. (1997) **The Architecture of the Language Faculty**. Cambridge, Massachusetts: MIT Press.

KILGARRIFF, A. (2000). "I don't believe in word senses". http://www.kcl.ac.uk/humanities/cch/seminar/99-00/seminar_kilgarriff.html

LAKOFF, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

MARTINS, H. (1999) "**Metáfora e Polissemia no estudo das línguas do mundo: uma aproximação não representacionista**". Tese de Doutorado inédita, UFRJ.

_____. (2003) "**Três Caminhos na Filosofia da Linguagem**" In Mussalim & Bentes (org.) *Introdução à Lingüística*. Volume III, São Paulo, Cortez Editora. 439-474.

NIETZSCHE, F. (1978 [1873]). "**Sobre verdade e mentira no sentido extra-moral**". Coleção *Os Pensadores*. São Paulo: Abril Cultural.

PINKER, S. (1995). *The Language Instinct*. New York: Harper Perennial.

SCHERER, M. (2002) "**Uma questão de vocabulário: considerações sobre o campo lexical no ensino de português para estrangeiros**". Dissertação de Mestrado inédita.

Milena Uzeda-Garrão

PUC-Rio.

SINCLAIR, J. (1991). **Corpus, Concordante, Collocation**. Oxford University Press, Oxford.

SOUZA FILHO, D. (1997) **Iniciação à história da Filosofia**. Rio de Janeiro: Jorge Zahar.

VERLINDE, S & SELVA, J. (2001) "**Corpus-based versus intuition-based lexicography: defining a word list for a French learners**" dictionary. In Proceedings of the Corpus Linguistics 2001 Conference. Lancaster University, UK. www.kuleuven.ac.be/ilt/grelep/publicat/verlinde.pdf.

WITTGENSTEIN, L. (1979 [1953]) **Investigações Filosóficas**. Coleção Os Pensadores, São Paulo: Abril Cultural.