



# **Portal multimodal/multilíngue para o avanço da ciência aberta nas Humanidades: Desenvolvimento da interface de acesso e banco de dados de metáforas linguísticas.**

Pontifícia Universidade Católica de São Paulo  
Programa: Linguística Aplicada e Estudos da Linguagem  
Doutorando: Marcos Roberto de Oliveira  
Orientador: Tony Berber Sardinha

O presente projeto de pesquisa está em consonância com o projeto institucional **'Portal multimodal/multilíngue para o avanço da ciência aberta nas Humanidades'** apresentado pelo PEPG em LAEL em resposta à chamada CNPq 25/2020. A construção do portal visa a promover a colaboração entre membros da comunidade científica e de fora dela por meio da reunião e disponibilização gratuita de dados, práticas metodológicas bem como a disseminação mais democrática do conhecimento.

## Objetivo principal

Verificar o impacto da reprodução de substantivos, adjetivos e verbos com usos metafóricos, e seus padrões colocacionais, no grau de naturalidade de textos que compõem o *corpus* de aprendiz (BR-ICLE).

## Base teórica

Os textos, que compõem o *corpus* de aprendiz (BR-ICLE), serão analisados sob a perspectiva da **teoria da metáfora conceptual** e da abordagem da **linguística de *corpus***.

## **Metáforas conceptuais e Linguística de corpus**

As vertentes mais tradicionais do estudo de metáforas, ou seja, aquelas que antecedem a publicação de *Metaphors We live by* por George Lakoff e Mark Johnson em 1980, tratam a metáfora como apenas figura de linguagem ou recurso retórico usado para comunicar ideias de forma mais eficiente ou para “embelezar” a linguagem em textos orais ou escritos.

*A vida dela é um conto de fadas.*

*A casa deles era uma prisão.*

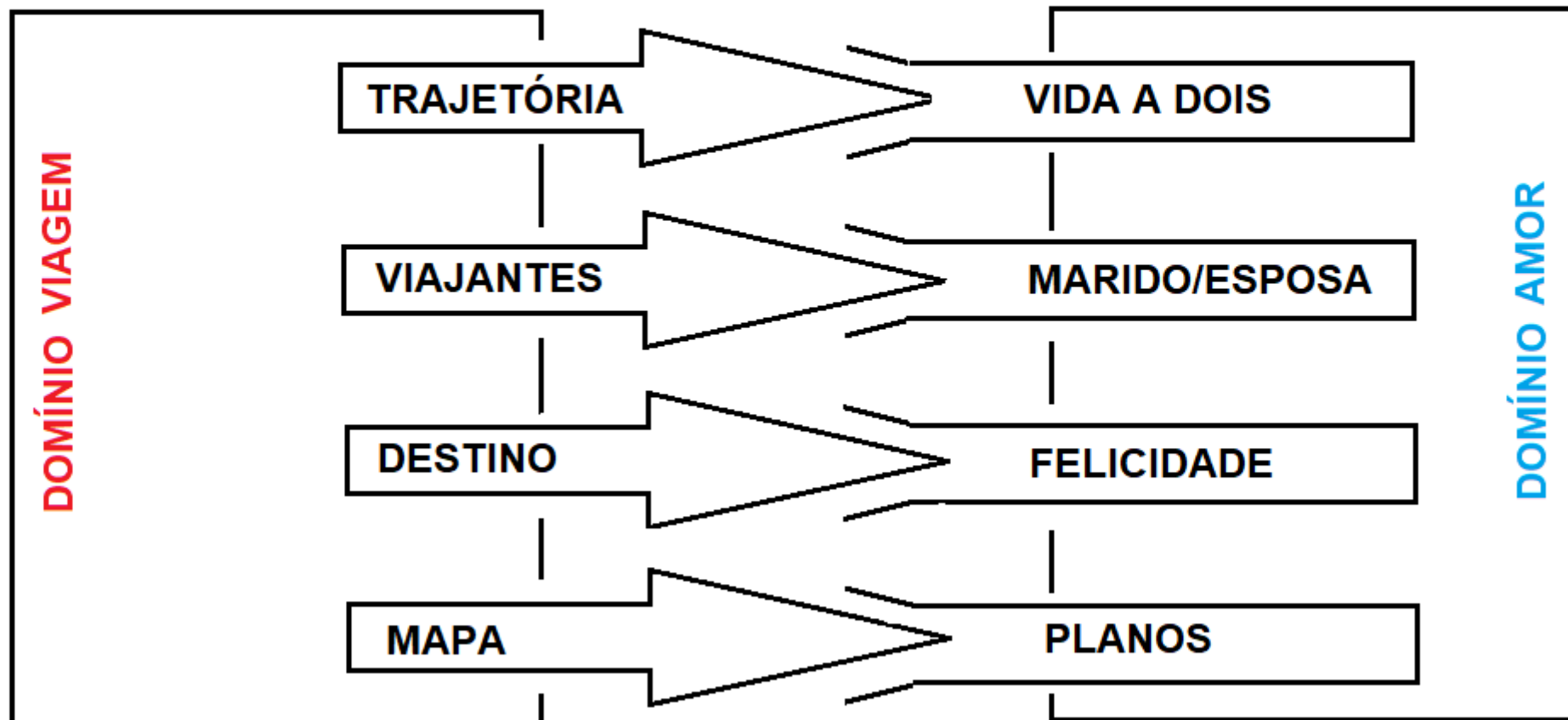
Já as metáforas conceptuais (LAKOFF e JOHNSON, 1980) são construtos mentais que antecedem sua manifestação em formas linguísticas. Elas consistem na correlação entre dois domínios conceptuais. O domínio fonte, que é um conceito mais concreto e que resulta de nossa experiência moto-sensorial com o mundo que nos cerca, e o domínio alvo, um conceito mais abstrato cuja compreensão é facilitada por elementos que constituem o domínio fonte.

A metáfora conceptual O AMOR É UMA VIAGEM, por exemplo, representa a conceptualização do domínio amor (conceito abstrato) por meio da projeção de elementos que constituem o domínio viagem (conceito concreto). Essa projeção ou mapeamentos (*mappings*) é possível porque alguns elementos que compõem o domínio viagem, encontram elementos correspondentes no domínio **amor**. Em outras palavras, nos baseamos em nosso conhecimento coerentemente organizado sobre o conceito **viagem** para assim conceptualizar o conceito de amor (GIBBS, 1999; LAKOFF e JOHNSON, 1980; BERBER SARDINHA 2007; KÖVECSES, 2010).



**CONCEITO CONCRETO**

**CONCEITO ABSTRATO**



Segundo Kövecses (2010), apesar de não termos consciência desses mapeamentos em nosso sistema conceptual, ele tem importante papel na definição de nossa realidade, pois estrutura tudo o que conceptualizamos, como pensamos e, principalmente, como nos comunicamos através de formas linguísticas.

Assim, considerando que a comunicação também é governada por esse sistema, a teoria nos leva ao entendimento de que a principal evidência da existência das metáforas conceptuais está na realização de metáforas linguísticas como:

*Nosso casamento está **indo** muito bem.*

*Esse relacionamento **chegou ao fim da linha**.*

***Nós trilhamos o caminho** do amor e respeito mútuo.*

As metáforas linguísticas são formadas por palavras ou ***associações de palavras*** (DEIGNAN, 2005) herdadas da linguagem ou terminologia usada na denominação de elementos de um conceito concreto (domínio fonte) que foram convencionalizadas por uma comunidade linguística para também nomear elementos correspondentes em um conceito abstrato (domínio alvo)

Para Kövecses (2010), as metáforas linguísticas também revelam que o conhecimento inconsciente que temos do domínio fonte vai além dos elementos envolvidos no mapeamento metafórico. Por exemplo, se O AMOR É UMA VIAGEM em nosso sistema conceptual, então o conhecimento de que uma viagem pode ser longa e cansativa pode nos levar à inferência ou **desdobramento metafórico (*entailments*)** de que um casal que vive junto há muito tempo pode cansar da relação (BERBER SARDINHA, 2007b).

## Exemplos de metáforas conceptuais e possíveis realizações linguísticas:

Metáfora conceptual	Metáfora linguística
ARGUMENTAÇÃO É GUERRA	<i>Ele usou todas as <b>armas</b> na <b>defesa</b> de seus argumentos</i>
TEORIAS SÃO EDIFICAÇÕES	<i>Chomsky <b>construiu</b> o <b>empreendimento</b> gerativista.</i>
IDEIA É COMIDA	<i>Ainda não consegui <b>digerir</b> aquilo que ele me disse!</i>
IDEAS SÃO OBJETOS	<i>Não consegui <b>pegar</b> nada do ele falou. Preciso melhorar meu francês.</i>
COMUNICAÇÃO É TROCA DE OBJETOS	<i>Meu orientador me <b>deu</b> algumas <b>sugestões</b> para o introdução do artigo.</i>

É importante lembrar que embora muitas metáforas conceptuais sejam universais, pois derivam de interações físicas experienciadas pela maioria dos humanos, há importantes **diferenças culturais** que determinam quais metáforas conceptuais são formadas e como seus elementos ou desdobramentos metafóricos são manifestados em metáforas linguísticas convencionadas por diferentes comunidades linguísticas (KÖVECSES, 2010). **Tal fato pode significar implicações importantes para aprendizes de língua estrangeira.**

O desconhecimento de como as metáforas conceptuais são materializadas por meio de palavras ou associações de palavras (colocações) por falantes nativos de língua inglesa em determinados registros, por exemplo, pode levar os aprendizes a se valerem do **princípio de livre escolha**, que segundo Sinclair (1991) opera quando o falante faz uso de palavras com significados fixos e as combinam de modo individual em função de um sistema mental de regras gramaticais, em detrimento do uso de colocações que, dentre outras vantagens, poderiam conferir maior naturalidade e fluência em seus textos falados e escritos.



## Hipótese

Baseando-se nos fatos apresentados até aqui e em estudos anteriores que demonstram que os substantivos, adjetivos e verbos são as categorias que mais frequentemente carregam significados metafóricos (GOATLY, 1997 e CAMERON, 2003), a hipótese formulada para este estudo é que a reprodução dessas categorias e de seus padrões colocacionais com usos metafóricos convencionalizados na variedade americana da língua inglesa, também contribuem para o grau de naturalidade da linguagem produzida por aprendizes de inglês.

A confirmação ou refutação dessa hipótese partirá da identificação dos substantivos, verbos e adjetivos com usos metafóricos, e seus padrões colocacionais mais frequentes nos textos do *corpus* BR-ICLE sob a perspectiva da teoria da metáfora conceptual e conceitos metodológicos da Linguística de corpus. Essa análise deverá responder as seguintes perguntas de pesquisa:

- 1) Quais são os substantivos, adjetivos e verbos com usos metafóricos mais frequentes nos textos do *corpus* BR-ICLE?
- 2) Quais são os colocados desses substantivos, adjetivos e verbos?
- 3) Quais são os padrões colocacionais formados por esses substantivos, adjetivos e verbos e seus colocados?
- 4) Qual é a frequência desses substantivos, adjetivos e verbos e seus colocados nos registros do *corpus* COCA?

## Linguística de corpus X Teoria da metáfora conceptual

Em um primeiro momento, o fato de a Linguística de *corpus* ter a linguagem utilizada em contextos reais de comunicação como principal objeto de estudo e a teoria da metáfora conceptual se ocupar do estudo de construtos que atuam no pensamento parece representar importante divergência teórica para o trabalho de pesquisa aqui proposto. Contudo, tendo em vista que a realização deste estudo propõe uma investigação de **linguagem em uso**, por meio da identificação das materializações linguísticas das metáforas conceptuais, a perspectiva da Linguística de *corpus* se faz coerente e necessária para a realização deste trabalho de pesquisa.

## **Metodologia**

Com o objetivo de responder as perguntas de pesquisa e, assim confirmar ou refutar a hipótese elaborada para o presente trabalho de pesquisa, seguiremos os passos descritos a seguir.

- 1) Identificação automática das categorias gramaticais presentes **no corpus de estudo** (BR-ICLE) com a kit de ferramentas computacionais The Stanford CoreNLP Natural Language Processing (MANNING, C. D. *et al.*, 2014). O BR-ICLE é um *subcorpora* que integra o *International corpus of Learner English* (ICLE) e sua versão mais recente (ICLEv3) é composta por 412 textos argumentativos de, em média, 500 palavras cada, escritos por estudantes de universidades brasileiras e aprendizes de língua inglesa de nível B2, segundo o Quadro Comum Europeu de Referência para Línguas (GRANGER, 2020).
- 2) Identificação dos dez substantivos, adjetivos e verbos mais frequentes ao longo do *corpus* BR- ICLE por meio de um *script* computacional em *Python*.

3) Criação das listas de colocados, ou seja, listas de palavras que ocorrem em torno dos substantivos, adjetivos e verbos com maior frequência de ocorrência ao longo do *corpus* . Essa tarefa será realizada por meio da ferramenta *Collocates*, disponível no *AntConc* 3.5.8 (ANTHONY, 2019), um kit de ferramentas de análise de textos que pode ser baixado gratuitamente.

4) Análise das linhas de concordância formadas pelos substantivos, adjetivos e verbos com o auxílio da ferramenta *Concordance*, também disponível no *AntConc* 3.5.8. para que os usos metafóricos sejam separados dos usos literais.

5) Verificação e análise da frequência de ocorrência dos padrões colocacionais formados pelos substantivos, adjetivos e verbos encontrados no corpus BR-ICLE no **corpus de referência**. O *corpus* referência selecionado para o estudo é o *Corpus of Contemporary American English (COCA)*. Um *corpus* multiregistro de monitoramento da variedade americana que integra a ferramenta online *English-Corpora.org* e atualmente é composto por mais de 1 bilhão de palavras balanceadas entre os registros ficção, revistas populares periódicos acadêmicos, conteúdo de websites, legendas de TV e filmes, blogs e transcrições de entrevistas de televisão e programas de rádio (DAVIES, 2020).



## REFERÊNCIAS BIBLIOGRÁFICAS

ANTHONY, L. **AntConc** (Version 3.5.8) Tokyo. Waseda University, 2019.

BERBER SARDINHA, T. A. **Linguística de *Corpus***. São Paulo: Manole, 2004.

BERBER SARDINHA, T. Análise de metáfora em *corpora*. **Ilha do Desterro**, V. 52, p. 167-1002, 2007a.

BERBER SARDINHA, T. **Metáfora**. São Paulo: Parábola, 2007b.

CAMERON, L. **Metaphor in Educational Discourse**. London: Continuum, 2003.

DANCYGIER, B. e SWEETSER, E. **Figurative language**. Cambridge: Cambridge University Press, 2014.

DAVIS, M. **Corpus of contemporary American English (COCA)**. Disponível em: <<https://www.english-corpora.org/coca>>. Acesso em: 10 de junho de 2020.

DEIGNAN, A. **Metaphor and *Corpus* linguistics**. Amsterdam/ Philadelphia: John Benjamins, 2005.

FECHER, B.; FRIESIKE, S. Open Science: One term, five schools of thought. In: BARTLING, Sönke; FRIESIKE, Sascha. (Eds). **Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing**. Springer, 2019, p. 17-48

FIRTH, J. **Papers in Linguistics**. Oxford, UK: Oxford University Press, 1957.

GOATLY, A. **The Language of Metaphors**. London: Routledge, 1997.

GRANGER, S *et al.* **international corpus of Learner English – Version 3**. Louvain-la-Neuve: UCL Presses Universitaires de Louvain, 2020.

KNÖCHELMANN, M. Open Science in the Humanities, or: Open Humanities? **Publications**, v. 7, n. 65, 2019.

KÖVECSES, Z. **Metaphor - A Practical Introduction**. 2. ed. New York: Oxford University Press, 2010.

LAKOFF, G. e JOHNSON, M. **Metaphors We Live By**. Chicago: University of Chicago Press, 1980.

LITTLEMORE J. e GRAHAM, L. **Figurative Thinking and Foreign Language Learning**. Basingstoke, U.K.: Palgrave Macmillan, 2006.

MANNING, Christopher D. *et al.* **The Stanford CoreNLP Natural Language Processing Toolkit** In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 2014, pp. 55-60.

OLIVEIRA, L. P. Linguística de *corpus*: Teoria, Interfaces e Aplicações. **Matraga**, Rio de Janeiro, v.16, n. 24, jan./jun. 2009.

REDDY, M. The Conduit Metaphor – a case of frame conflict in our language about language. In: Andrew Ortony (ed.), *Metaphor and thought*. Cambridge University Press, 1979.

ROSEN, D. E.; PURINTON, E. **Website design**: Viewing the web as a cognitive landscape. *Journal of Business Research*, v. 57, n. 7, p. 787-794, 2004.

SILVA, F.C.C e SILVEIRA, L. O ecossistema da Ciência aberta. **Transformação**, v. 31, 2019. Disponível em: <<http://www.ref.scielo.org/kmkdcc>> Acesso em:10 de junho de 2021.

SINCLAIR, J. **Corpus, concordance, collocation**. Oxford: Oxford University Press, 1991.

SINCLAIR, J. Naturalness in Language. In: **Corpus Linguistics Aarts J. and Meys W.** (eds.), 1983, p. 203-10.

STEEN, G, *et al.* **A Method for Linguistic Metaphor identification: From MIP to MIPVU**. Amsterdam: John Benjamins, 2010.

TOGNINI-BONELLI, E. **Corpus Linguistics at Work**. Amsterdam: John Benjamins, 2001.