

USO DE UNIDADES FRASEOLÓGICAS NA PRODUÇÃO ESCRITA EM PORTUGUÊS COMO LÍNGUA ADICIONAL: ANÁLISE DE COLIGAÇÕES E COLOCAÇÕES LEXICAIS

GRACIELE CORDEIRO¹; ALESSANDRA BALDO²

¹ Centro Universitário Internacional – grahcord@gmail.com (autora)

² Universidade Federal de Pelotas – alessabaldo@gmail.com (orientadora)

1. INTRODUÇÃO

O estudo da linguagem tem expandido seu escopo para incluir as Unidades Fraseológicas (UFs), que desempenham papel crucial na fluência e na naturalidade da comunicação (MATIAS, 2008). A partir da segunda metade do século XX em diante, com o desenvolvimento dos estudos pragmáticos, discursivos e cognitivos da linguagem, para citar alguns dos mais relevantes, a linguagem figurada deixou de ser vista como mero adorno estilístico para ser reconhecida como componente intrínseco do pensamento e da expressão humana (GEDOZ, 2016), elemento constitutivo da forma como concebemos o mundo, configurando-se como reflexo de processos cognitivos, sociais e culturais (SILVA; ISQUERDO, 2020).

A fraseologia emerge como subdisciplina da lexicologia, dedicada ao estudo científico das UFs. Embora a definição de "fraseologia" seja complexa e não haja conceito abrangente único, ela é entendida tanto como a disciplina que estuda as UFs quanto como o próprio conjunto dessas unidades. As UFs constituem parcela significativa do repertório linguístico, representando entre 30% e 50% do discurso fluente (SINCLAIR, 1991; ERMAN; WARREN, 2000), e são definidas como sequências de palavras ou outros elementos que, por sua natureza pré-fabricada, são armazenadas e recuperadas como um todo na memória (WRAY, 2002). Ainda, são caracterizadas por três traços fundamentais: a polilexicalidade, sua composição por múltiplos elementos lexicais; a fixidez ou estabilidade, manifestada em diferentes graus de cristalização estrutural; e a idiomaticidade, expressa na não-composicionalidade semântica (TAGNIN, 2013). Estes traços distribuem-se em contínuos complexos que vão desde colocações relativamente transparentes até expressões idiomáticas completamente opacas.

Nesse sentido, cabe relembrar a definição de "falante ingênuo" proposta por Fillmore (1979), caracterizado como aquele que se encontra em desvantagem comunicativa por desconhecer as convenções fraseológicas de uma língua, independentemente de seu domínio das regras gramaticais. Tal limitação assume particular relevância no contexto do ensino de Português como Língua Adicional (PLA), posto que a competência fraseológica transcende a dimensão puramente linguística.

A complexidade inerente aos traços das UFs manifesta-se em dificuldades particulares na sua aquisição por aprendizes de línguas estrangeiras. Neste sentido, o presente trabalho propõe-se a examinar o papel das UFs no ensino-aprendizagem de PLA. O objetivo principal é identificar as características e principais dificuldades nos usos de combinações lexicais, em especial as coligações e as colocações, por aprendizes de PLA de nível intermediário e avançado em tarefas de produção escrita.

No que diz respeito às colocações lexicais, os objetivos específicos são: (i) compreender em que medida o uso das colocações lexicais pelos aprendizes é influenciado por sua língua materna; (ii) mapear as estruturas cujo emprego das colocações se apresente mais difícil aos aprendizes, e (iii) propor atividades didáticas que auxiliem na superação dessas dificuldades. Em relação às coligações lexicais, os objetivos incluem: (iv) identificar em que medida a utilização das coligações é condicionada pela língua materna; (v) verificar a frequência relativa de determinadas coligações em português em comparação a outras presentes nas produções escritas; (vi) reconhecer padrões recorrentes nas dificuldades de uso das coligações; (vii) elaborar atividades didáticas voltadas à superação das dificuldades.

À luz destas considerações, o presente estudo articula perspectivas teóricas contemporâneas com análises empíricas. A investigação sistemática das UFs no contexto de PLA ainda apresenta lacunas significativas. Pesquisas com *corpora* de aprendizes têm revelado padrões persistentes de desvios colocacionais (ALMEIDA, 2014; ROZOVSKAYA; ROTH; SRIKUMAR, 2014) e de transferência interlinguística inadequada (TSCHICHOLD, 2003). As dificuldades documentadas no seu uso evidenciam que a exposição limitada ao *input* linguístico autêntico compromete os processos naturais de consolidação através da frequência, conforme apontado por Granger (1998) e Wray (2002). A natureza multifacetada do fenômeno, aliada aos desafios de aquisição de línguas, exige investigações que informem as práticas de ensino. É nessa lacuna didática que este trabalho se enquadra.

2. METODOLOGIA

A metodologia é de natureza qualitativo-quantitativa, com base em uma análise de *corpus*. O *corpus* é constituído por cem produções escritas de aprendizes adultos de PLA, com níveis de proficiência intermediário e avançado, que participaram de cursos preparatórios para o exame CELPE-Bras, a certificação oficial exigida para comprovação de proficiência em português. O corpus foi obtido a partir de produções das tarefas 3 e 4 do referido exame.

Em uma primeira etapa, são identificadas as expressões lexicais cujo uso diverge daquele esperado no português brasileiro em situações de produção textual nos gêneros em análise. Essas expressões são etiquetadas segundo a taxonomia das UFs proposta por Tagnin (2013). Para cada UF identificada, os dados serão tabulados em tabelas específicas, permitindo uma análise estatística dos padrões de erro, bem como uma análise qualitativa dos fatores responsáveis pela dissonância. Busca-se compreender os fatores responsáveis por tais divergências, considerando três aspectos principais: o grau de similaridade da unidade lexical em relação à língua materna do aprendiz, a adequação ou inadequação do uso conforme o contexto comunicativo e a ocorrência de substituição por outra expressão próxima em português.

O estudo ainda prevê uma segunda fase, isto é, a elaboração de atividades didáticas, que será baseada nos princípios teórico-metodológicos da criação de sequências-didáticas. Aqui, elas são entendidas como um conjunto de atividades escolares organizadas, de maneira sistemática, em torno de um gênero textual oral ou escrito, com o objetivo de dar acesso aos alunos a práticas de linguagem novas ou dificilmente domináveis (DOLZ; NOVERRAZ; SCHNEUWLY, 2004).

Como essa fase do estudo ainda não foi iniciada, não nos deteremos na descrição de seus procedimentos metodológicos neste momento.

3. RESULTADOS E DISCUSSÃO

A análise preliminar do *corpus* revelou padrões recorrentes de desvios colocacionais e de uso de expressões conectivas que evidenciam a transferência interlinguística negativa. Foram identificadas diversas inadequações, particularmente por parte de falantes de espanhol. Exemplos incluem o uso de “bagagem DA mão” (em vez de “bagagem DE mão”), “fazer um tempo” (por “reservar/deixar/separar um tempo”) e “igualdade sexual” (“igualdade de gênero”). Nas expressões conectivas, destacam-se exemplos como “no mesmo tempo” (em vez de “ao mesmo tempo”) e “pelo que” (“por isso”). Esses desvios comprometem a idiomatidade e a coesão do texto.

Tais resultados reforçam a constatação de que o falante “ingênuo” (FILLMORE, 1979) tende a traduzir literalmente combinações de sua L1, o que resulta em construções não idiomáticas no PLA. Nesses casos, a proximidade formal entre línguas pode mascarar diferenças fundamentais nas convenções fraseológicas. O uso de verbos de alta frequência em detrimento de alternativas mais específicas (“fazer” em vez de “reservar”) também aponta para uma estratégia compensatória que, embora funcional, compromete a fluência nativa.

As observações preliminares do *corpus* estão em consonância com pesquisas anteriores (ROZOVSKAYA; ROTH; SRIKUMAR, 2014) que indicam que erros na seleção verbal constituem a categoria mais problemática. A dificuldade de domínio de colocações verbais deslexicalizadas configura-se como uma barreira significativa para o alcance da proficiência nativa em PLA. A análise dos desvios demonstra que a aquisição de UFs é um processo complexo e não-linear, em que os mecanismos de *chunking* e de processamento holístico podem ser responsáveis pelo surgimento de padrões de erro persistentes na interlíngua dos aprendizes, levando à fossilização de estruturas não-idiomáticas.

4. CONCLUSÕES

A presente pesquisa contribui de forma significativa para a área de ensino de PLA, pois evidencia a centralidade do componente fraseológico para a competência comunicativa e mapeia as principais dificuldades enfrentadas pelos aprendizes com relação ao uso adequado das UFs. A principal inovação do trabalho reside em sua natureza aplicada, apontando para a necessidade de abordagens de ensino mais eficazes e teoricamente informadas que transcendam a mera apresentação de listas de expressões.

Os resultados parciais demonstram a importância do ensino de UFs ser sistemático, contextualizado e baseado em evidências de *corpus*. A pesquisa busca oferecer diretrizes pedagógicas para auxiliar os alunos a superarem os desvios e a fossilização, contribuindo para a internalização dessas unidades como componentes essenciais de seu repertório ativo. O estudo reafirma que o conhecimento de sequências formulaicas aprimora substancialmente a coesão textual, configurando-se como marcador distintivo de proficiência avançada.

Ainda que seja precoce especular sobre as causas das dificuldades observadas nos textos a serem analisados, pode-se antecipar que os resultados deste estudo deverão ter duas principais repercussões. A primeira, de caráter

didático, é fornecer informações concretas para a elaboração de materiais pedagógicos mais eficazes. A segunda, de natureza teórica, é a de poder contribuir para o debate sobre os processos de aquisição do léxico bilíngue, especialmente no que diz respeito às relações de inferência e interferência entre a língua materna e o português como L2 dos aprendizes (CORDER, 1982). O presente trabalho, ao articular perspectivas teóricas com análises empiricamente fundamentadas, busca não apenas contribuir para a consolidação de uma base científica sólida, mas também catalisar o desenvolvimento de recursos pedagógicos e futuras investigações que possam efetivamente capacitar aprendizes de PLA a transcenderem as limitações do falante "ingênuo".

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, V. C. **Investigando colocações em um corpus de aprendiz**. 2014. Tese (Doutorado em Linguística Aplicada). Programa de Pós-Graduação em Linguística Aplicada, Universidade Federal de Minas Gerais.
- CORDER, S. P. **Error Analysis, Interlanguage and Second Language Acquisition**. *Language Teaching & Linguistics: Abstracts*, v. 8, n. 4, 201-218, 1982.
- DOLZ, J.; NOVERRAZ, M.; SCHNEUWLY, B. Sequências didáticas para o oral e a escrita: apresentação de um procedimento. In: SCHNEUWLY, B.; DOLZ, J. (orgs.) **Gêneros orais e escritos na escola**. Campinas: Mercado de Letras, 2004.
- ERMAN, B.; WARREN, B. The idiom principle and the open choice principle. *Text*, v. 20, n. 1, p. 63-76, 2000.
- FILLMORE, C. J. Innocence: A second idealization for linguistics. **Proceedings of the Annual Meeting of the Berkeley Linguistics Society**, v. 5, p. 63-63, 1979.
- GEDOZ, S. Análise linguística e reescrita textual: articulando encaminhamentos. **Fórum Linguístico**, v. 13, n. 2, p. 1225-1239, 2016.
- MATIAS, L. C. **Expressões idiomáticas corporais no Dicionário Bilingüe de Uso Espanhol-Português / Português-Espanhol (DiBu)**. 2008. Dissertação (Mestrado em Estudos da Tradução). Programa de Pós-Graduação em Estudos da Tradução, Universidade Federal de Santa Catarina.
- ROZOVSKAYA, A.; ROTH, D.; SRIKUMAR, V. Correcting grammatical verb errors. **Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics**, p. 358-367, 2014.
- SILVA, C. A. N.; ISQUERDO, A. N. Fraseo (topônimos): Um estudo de topônimos polilexicais na perspectiva da fraseologia. **Revista do GEL**, v. 17, p. 286-308, 2020.
- SINCLAIR, J. **Corpus Concordance and Collocation**. Oxford University Press, 1991.
- TAGNIN, E. **O jeito que a gente diz: expressões consagradas em inglês e português**. DISAL, 2013.
- TSCHICHOLD, C. **English core linguistics: essays in honour of D.J. Allerton**. Peter Lang, 2003.
- WRAY, A. **Formulaic Language and the Lexicon**. Cambridge University Press, 2002.