

AVALIAÇÃO AUTOMÁTICA DE TEXTOS EM FRANCÊS

Giulia Victoria Proença Felix (PIBIC/CNPq/FA/UEM), Juliano Desiderato Antonio (Orientador). E-mail: jdantonio@uem.br.

Universidade Estadual de Maringá, Centro de Ciências Humanas, Letras e Artes, Maringá, PR.

Linguística, Letras e Artes/ Teoria e Análise Linguística

Palavras-chave: complexidade textual; FLE; avaliação automática.

RESUMO

O objetivo principal deste trabalho é utilizar um analisador automático de complexidade textual para compilar uma seleção de textos de diversos níveis que possam ser utilizados por professores de FLE. A ferramenta utilizada é o DMesure, desenvolvido para a classificação automática dos textos de acordo com os níveis do Quadro Europeu Comum de Referência para as Línguas. A avaliação automática dos textos foi verificada por três experientes professores de FLE. Em termos de resultados, observou-se que, na maioria dos casos, os professores validaram a avaliação realizada pela ferramenta.

INTRODUÇÃO

Com as recentes evoluções no campo do Processamento de Linguagem Natural (PLN), cada vez mais tarefas linguisticamente complexas vêm sendo realizadas por ferramentas computacionais. Alguns exemplos são o reconhecimento da fala humana, a tradução automática, a sumarização automática etc. Uma outra possibilidade interessante que vem se desenvolvendo é a possibilidade de se analisar automaticamente a complexidade de textos do ponto de vista linguístico. O objetivo principal deste projeto é utilizar um analisador automático de complexidade textual para criar uma lista de textos de diversos níveis que possam ser utilizados por professores de FLE. A ferramenta utilizada é o DMesure, disponível em <<https://cental.uclouvain.be/dmeasure>>, desenvolvido por Yancey, Pintard e François (2021) para a classificação automática dos textos de acordo com os níveis do Quadro Europeu Comum de Referência para as Línguas (doravante QECRL). A

classificação tem cinco níveis de dificuldade: A1 (mais fácil), A2, B1, B2 e C1 (mais difícil).

MATERIAIS E MÉTODOS

A pesquisa foi realizada em três etapas. Inicialmente, selecionaram-se trinta textos em francês de diversos níveis e de diversos gêneros.

Na segunda etapa, os textos foram processados pela ferramenta computacional DMesure para se verificar seu nível no que se refere ao QECRL. A ferramenta analisa a leitura dos textos. Ponomarenko e Evers (2022, p. 42) definem leitura como “potencial facilidade ou dificuldade de leitura de um texto”, levando em conta não apenas fatores linguísticos mas também o perfil do leitor pretendido pelo texto. De acordo com Yasseri, Kornai e Kertész (2012), a leitura é um dos principais temas relacionados à complexidade linguística.

Na terceira e última etapa, solicitou-se a três experientes professores de Francês como Língua Estrangeira (FLE) que validassem ou não a classificação apresentada pela ferramenta computacional. No caso de não concordância com a análise automática, solicitou-se que justificassem o motivo da discrepância.

RESULTADOS E DISCUSSÃO

Na tabela 1 a seguir, apresenta-se a frequência de concordância entre os professores que validaram a classificação automática dos textos realizada por meio da ferramenta DMesure.

Tabela 1. Concordância dos validadores

	N	%
Todos os avaliadores concordam com a classificação automática	8	26,67
Dois avaliadores concordam com a classificação automática	14	46,67
Um avaliador concorda	6	20

com a classificação automática		
Nenhum avaliador concorda com a classificação automática	2	6,67
Total	30	100

Como pode ser observado, a maior frequência de concordância com a classificação da ferramenta é entre dois validadores (46,67%). Em segundo lugar, aparece a concordância dos três validadores com a ferramenta (26,67%). Somando-se esses dois resultados, obtém-se 73,34%, ou seja, trata-se de um resultado favorável para a ferramenta automática, uma vez que, na maioria dos textos, dois ou três validadores concordaram com a classificação da plataforma.

A discordância dos três validadores com a ferramenta ocorreu em apenas 6,67% dos casos e a concordância de apenas um validador em 20%. A discrepância entre as classificações se deu, principalmente, em função de sofisticação lexical, uso de determinados tempos verbais e de algumas estruturas gramaticais. Tome-se como exemplo um dos textos que compõem o *corpus* de análise do trabalho. No excerto de “L’oiseau blanc”, de Alex Cousseau, classificado como A2/B1 pela ferramenta, dois validadores concordaram com a classificação. O professor que discordou considerou que o texto deveria ser classificado apenas como A2 pelo fato de a linguagem ser simples e sem vocabulário complexo.

“Le jour où commence cette histoire,
le vent souffle fort, plus fort que d’habitude.
Un océan sépare la terre de l’Amérique et la terre de l’Europe.
De chaque côté, des hommes vivent sans se connaître.
Mais le vent s’en moque.
Le vent souffle pareil ici et là.

Et par son souffle, les rivières gonflent et chantent plus vivement,
les herbes se couchent, la poussière s’envole des chemins.
Des vagues se forment, des nuages se dispersent.
C’est comme si le vent s’amusait à tout mettre
en désordre de chaque côté de l’horizon.”

CONCLUSÕES

Para o professor, trata-se de uma grande vantagem utilizar uma ferramenta como o DMeasure, uma vez que a aplicação computacional pode auxiliá-lo a decidir se um determinado texto é pertinente no processo de ensino-aprendizagem e adequado às capacidades de seus alunos. Nos estudos em FLE, a escolha de materiais autênticos é fundamental para a eficácia do ensino da língua, conforme Courtillon (2003), isto significa buscar originalidade para compreender a realidade da linguagem. A realização desta permitiu verificar que, na maioria dos casos, as classificações realizadas pela plataforma automática foram validadas pelos professores.

AGRADECIMENTOS

Agradeço ao CNPq pelo apoio à realização deste projeto por meio da concessão de bolsa de Iniciação Científica.

REFERÊNCIAS

COURTILLON, J. **Élaborer un cours de FLE**. Paris: Hachette Livre, 2003.

MCMAMARA, D. S.; GRAESSER, A. C.; MCCARTHY, P. M.; CAI, Z. **Automated Evaluation of Text and Discourse with Coh-Metrix**. Cambridge: Cambridge University Press, 2014.

PONOMARENKO, G. L.; EVERS, A. Leiturabilidade e ensino: autores-base e seus trabalhos. In: FINATTO, M. J. B.; PARAGUASSU, L. B. (Orgs.) **Acessibilidade textual e terminológica**. Uberlândia: Edufu, 2022. p. 41-71.

YANCEY, K.; PINTARD, A.; FRANÇOIS, T. Investigating readability of French as a foreign language with deep learning and cognitive and pedagogical features. **Lingue e Linguaggio**, v. 2021, n. 2, p. 229–258, 2021. Acesso em: 27 agosto 2024.

YASSERI, T.; KORNAL, A.; KERTÉSZ, J. A Practical Approach to Language Complexity: A Wikipedia Case Study. **PLoS ONE**, v. 7, n. 11, p. e48386, 2012. Acesso em: 27 agosto 2024.