

ANÁLISE COMPARATIVA DA FREQUÊNCIA DAS PALAVRAS POR MEIO DA LINGUÍSTICA DE CORPUS E DA LINGUÍSTICA COMPUTACIONAL

Resumo

O objetivo deste estudo foi analisar as respostas dos educandos de um curso superior da Universidade Federal de Santa Maria, quando perguntado a eles num Fórum de Discussão na disciplina Educação Popular, Movimentos Sociais e Educação do Campo, ofertada no primeiro semestre de 2018, sobre como a Educação Popular pode contribuir para a prática pedagógica de um futuro educador. Para a efetivação do objetivo proposto por este estudo, foi realizado um estudo bibliográfico, de caráter descritivo e exploratório, com abordagem qualitativa e para a investigação das respostas obtidas no Fórum, aplicou-se uma análise comparativa da frequência das palavras, trata-se de uma exploração de dados linguísticos textuais com base na Linguística de Corpus e na Linguística Computacional. Evidenciou-se que o método é eficiente, refletindo por meio das palavras evidenciadas que o educando acredita que a Educação Popular contribui para a sua prática pedagógica.

Palavras-chave: Educação do Campo; Educação Popular; Fórum; AVEA; RStudio.

1. Introdução

O presente estudo contempla a análise das respostas dos educandos do curso superior de Licenciatura em Educação do Campo pela Universidade Federal de Santa Maria/UFSM, contemplado pela Universidade Aberta do Brasil (UAB), quando perguntado a eles sobre o que a Educação Popular pode contribuir para a prática pedagógica de um futuro educador. A pergunta foi feita em um Fórum de Discussão disposto no Ambiente Virtual de Aprendizagem (AVEA) Moodle, na disciplina de Educação Popular, Movimentos Sociais e Educação do Campo, ofertada no primeiro semestre de 2018.

Para a análise das respostas obtidas no Fórum aplicou-se uma análise comparativa da frequência das palavras, trata-se de uma exploração de dados linguísticos textuais com base na Linguística de Corpus e na Linguística Computacional. Conforme Kader e Richter (2013, p.13) a Linguística Corpus é:

Um campo que se dedica à criação e análise de corpora (plural latim de corpus), ou seja, conjunto de textos e transcrições de fala armazenadas em arquivos de computador. A Linguística de Corpus vem mudando a maneira como se investiga a linguagem, nos seus mais diversos níveis, colocando à disposição do analista quantidades de dados antes inacessíveis.

Assim, a Linguística de Corpus depende de programas de computador para lidar com a corpora. Dentre os softwares que existem para auxiliar a Linguística de Corpus foi escolhido a Linguística Computacional R (Versão 3.1.2). Segundo Oushiro (2014, p. 134) a Linguística Computacional R é uma linguagem de programação para a análise de dados “que pode ser utilizada para realizar computações estatísticas e gráficas, compilar e anotar corpora, produzir listas de frequências”, entre outras tarefas.

Porém, a Linguística Computacional R, tem um ambiente desenvolvido e integrado para o R que é o software livre RStudio (Versão 0.98.1087), e que disponibiliza “ferramentas

adicionais diretamente na interface gráfica, como a visualização dos scripts abertos recentemente, o histórico de linhas de comando executadas e a lista de pacotes instalados” (OUSHIRO, 2014, p. 136).

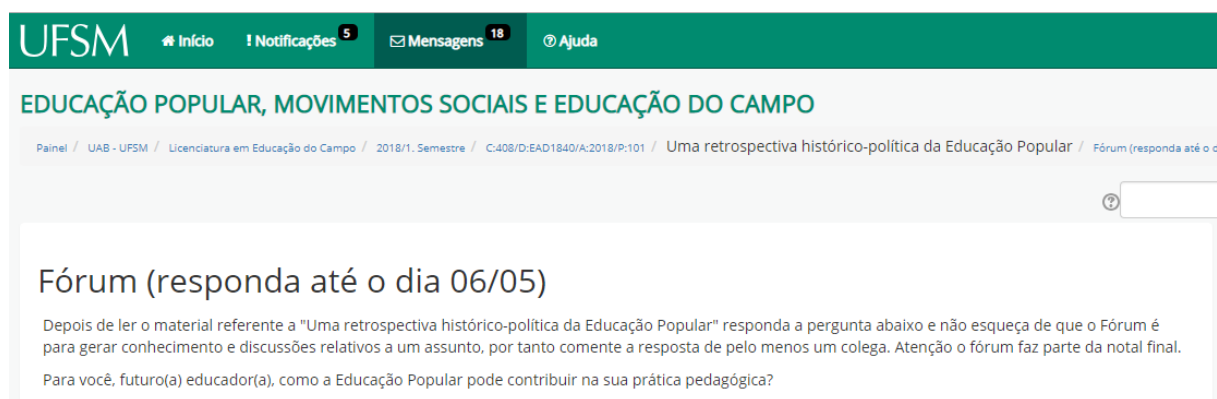
Um dos recursos para a visualização dos resultados de medição de frequência obtidos por meio desses programas é a nuvem de palavras (wordcloud), que consiste de um alinhamento de palavras representadas em tamanhos proporcionais à frequência com que aparecem em um conjunto de textos. Esse recurso será usado para gerar os resultados desta pesquisa, assim, este estudo objetiva utilizar um corpus significativo para demonstrar de que forma o software livre RStudio e a linguagem R, como recursos da Linguística Computacional, podem contribuir para a análise linguística baseada em Corpus, a fim de contemplar a análise das respostas dos educandos no Fórum de Discussão proposto a eles e se os usos das tecnologias digitais de informação e comunicação (TDICs) foram significativas para a reflexão, interatividade e interação dos educandos no Ambiente Virtual de Ensino Aprendizagem (AVEA) – Moodle.

2. Materiais e métodos

Para a efetivação do objetivo proposto por este estudo, foi realizado um levantamento bibliográfico, de caráter descritivo e exploratório, com abordagem qualitativa. Foi utilizado como população para a compilação dos dados obtidos, por meio de Fórum de Discussão proposto, os 150 educandos na disciplina de Educação Popular, Movimentos Sociais e Educação do Campo, do Curso de Licenciatura em Educação do Campo oferecido pela UFSM contemplado pela Universidade Aberta do Brasil (UAB), oferecido em cinco cidades do Rio Grande do Sul (RS), que estão nos polos de: Agudo, Cerro Largo, Itaqui, São Sepé e Seberi.

A atividade no Fórum de Discussão, disposto no AVEA Moodle, consistia em os educandos ler o material referente a "Uma retrospectiva histórico-política da Educação Popular" e responder uma pergunta, e ainda comentar a resposta de pelo menos um colega. A pergunta era: “Para você, futuro(a) educador(a), como a Educação Popular pode contribuir na sua prática pedagógica?”, conforme Figura 1 a seguir.

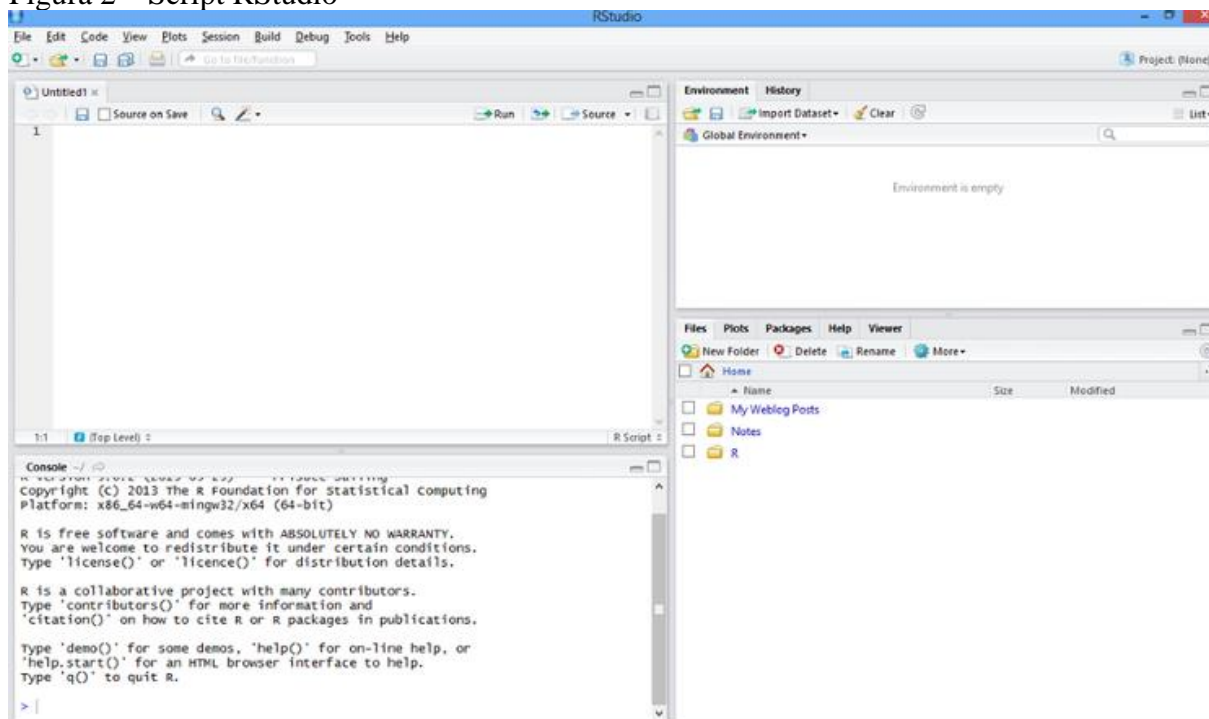
Figura 1 – Fórum de discussão do AVEA Moodle



Fonte: AVEA Moodle - UFSM, 2018.

Para tanto, a análise deu-se da seguinte forma: iniciou-se a interpretação do script no RStudio, conforme Figura 2, realizando inicialmente a limpeza da memória, a configuração do local de salvamento dos dados e o diretório de trabalho. Para a preparação do corpus, procedeu-se com: carregamento do pacote tm (Text Mining) no R; leitura dos arquivos do diretório; carregamento dos arquivos e criação do corpus; sumário dos dados; remoção dos marcadores e espaços em branco; análise das propriedades do corpus; alteração das letras maiúsculas para minúsculas; remoção das stopwords – palavras a serem ignoradas na busca (OUSHIRO, 2014), remoção da pontuação.

Figura 2 – Script RStudio



Fonte: Arquivo da autora, 2018.

Preparado o corpus, procedeu-se a matriz de frequência (criação, operação e associação dos termos). Feito isso, usou-se expressões regulares para carregar os pacotes Wordcloud, XML e RColorBrewer, usados na formação das nuvens de palavras. Por fim, foram inseridas as linhas de comando para a criação das nuvens com os seguintes argumentos e valores: escala (8, 4), frequência mínima (10), número máximo de palavras (Inf.), ordem randômica (False), proporção de palavras em rotação (15). Tais comandos foram executados nos trechos e parágrafos coletados do Fórum, resultando na formação de uma (1) nuvem de palavras.

3. Resultados e discussão

Buscando analisar as respostas dispostas no Fórum de Discussão, por meio da pergunta: “Para você, futuro educador, como a Educação Popular pode contribuir na sua prática pedagógica?”, nesta perspectiva, conforme Figura 3, pode-se verificar que o termo mais citado no Fórum é “desafio”, seguido de “realidade” e “conhecimento”.

É notório que em todos os trechos citados aparecem a palavra “desafio” e sempre está relacionado com as palavras “realidade” e/ou “conhecimento”. Por fim, esta análise ampla do corpus, dos trechos descritos pelos educandos no Fórum de Discussão proposto na disciplina Educação Popular, Movimentos Sociais e Educação do Campo, demonstram a importância do educador/tutor de trabalhar com as TDICs no contexto da EaD, a fim de proporcionar aos educandos reflexões, criticidade, interatividade e interação.

4. Conclusão

O uso da linguística computacional neste estudo, por meio da linguagem R e do software livre RStudio, favoreceu o exame dos trechos analisados postados no Fórum de Discussão pelos educadores e demonstrou, por meio do Word Cloud, as palavras mais citadas nos trabalhos científicos analisados, constituindo um panorama importante das informações existentes no material coletado. Da mesma forma, a pesquisa evidenciou que as palavras ressaltadas refletem o universo de como o educando, futuro educador, enxerga como a Educação Popular pode contribuir na sua prática pedagógica.

Logo, a educação voltada ao diálogo, para a reflexão, com troca de saberes entre o educador e o educando, buscando a construção de novos conhecimentos, proporciona ao educando uma educação para o pensar, refletir e criticar. É preciso que educadores desprendam da ideia de uma educação tradicional, onde o educando é apenas um depósito de conhecimentos, não possuidor do direito de questionar ou de expor seus conhecimentos prévios.

Futuros educadores precisam pensar e buscar uma formação de sujeitos com conhecimento e consciência cidadã e a organização do trabalho político para afirmação do sujeito, fazendo que o processo ensino e aprendizagem seja visto como ato de conhecimento e transformação social, incentivando-os a conhecerem a sociedade em que vivem, afirmando e validando suas culturas populares.

5. Referências

KADER, C. C. C.; RICHTER, M. G. Linguística de Corpus: Possibilidade e avanços. Instrumento: **R. Est. Pesq. Educ.** Juiz de Fora, v. 15, n. 1, jan./jun. 2013.

OUSHIRO, Livia. Tratamento de dados com o R para análises sociolinguísticas. In: FREITAG, Raquel Meister Ko. Freitag (Org.). **Metodologia de coleta e manipulação de dados em Sociolinguística**. São Paulo: Edgard Blücher, 2014. p. 133-176.

WERTHEIN, J. (org.) Educação de Adultos na América Latina. Campinas/SP: Papirus, 1985.