

Dissertation & OPEN ACCESS

DISTRIBUTIONAL CORPUS ANALYSIS OF KOREAN NEOLOGISMS USING ARTIFICIAL INTELLIGENCE

Wonbin Kim

Doctor of Philosophy (PHD), University of Georgia; 12/2022

Abstract

artificial intelligence | collocation analysis | distributional corpus analysis | distributional frequency profiles | Korean neologisms | semantic change

Distributional corpus analysis (DCA) is an approach which reveals lexical relations using large-scale corpora and computational techniques in natural language processing. It has an advantage of processing and analyzing lexical relations in a quantitative, consistent, and objective way. Although the DCA approach allows analysts to process large-scale linguistic data efficiently, there are few studies using the DCA approach to investigate language phenomena within corpus linguistics. Therefore, this study aims to bridge the gap between the DCA approach and corpus linguistics by designing and describing DCA from the perspective of corpus linguistics. Specifically, this study uses the DCA approach to analyze the distributional behaviors of three Korean neologisms leyal, lwuce, and kay- and track semantic change of the three neologisms. For the analysis of distributional behaviors, Korean Twitter data spanning about ten years is collected and three state-of-the-art techniques are employed. For leyal, word2vec and cosine similarity are used and for lwuce, Latent Dirichlet Allocation is employed. For kay-, long short-term memory is utilized.

Files and links



Dissertation_Wonbin Kim



Open Access

Metrics

Details

TITLE DISTRIBUTIONAL CORPUS ANALYSIS OF KOREAN NEOLOGISMS USING ARTIFICIAL INTELLIGENCE

CREATORS Wonbin Kim

CONTRIBUTORS William A. Kretzschmar (Advisor)

Linda Harklau (Committee Member) Paula J. Mellom (Committee Member)

AWARDING University of Georgia; Doctor of Philosophy (PHD) INSTITUTION