This thesis explores the assessment of word importance, from defining the concept to creating and evaluating a prediction system. We collect word importance labels using our web-based annotation tool and define word importance as word rankings. We propose a self-supervised machine learning method where new words are artificially inserted into text, and then we fine-tune the BERT model to learn to identify these words. We hypothesize that the resulting model will assign a higher likelihood of insertion to less important words. We experiment with two different insertion strategies: the List Inserting Method and the BERT Inserting Method. Evaluations on our collected data show that our methods outperform traditional baselines such as TF-IDF and rival existing approaches, demonstrating the effectiveness of our approach in predicting word importance.