

*Abstract*—Building a phonemic model of words is a description of a function in mapping the spelling of words to a series of phonetic symbols that represent the pronunciation of the word. Basically, the word phonemicization is the same as a grapheme-to-phoneme (G2P) conversion. The speech synthesis system from text needs to be done in the process of converting graphemes into phonemes, so a G2P is needed here. The conversion of graphemes into phonemes will represent the mapping of each grapheme or spelling symbol in any word to the phonemic representation or pronunciation symbol. In this paper, a word phonemicization model is developed using a bidirectional long short-term memory (BLSTM). A rule-based re-optimization procedure is proposed to enhance the model. An evaluation of a dataset of 50 k Indonesian formal words shows that the proposed model gives a phoneme error rate (PER) of 0.73%, which is much lower than all previous models. The errors are mostly caused by converting grapheme <e>.

*Keywords*—bidirectional long short-term memory, Indonesian phonemicization model, n-Gram