

Combined Automatic Speech Recognition and Machine Translation in Business Correspondence Domain for English-Croatian

Authors : Sanja Seljan, Ivan Dunder

Abstract : The paper presents combined automatic speech recognition (ASR) for English and machine translation (MT) for English and Croatian in the domain of business correspondence. The first part presents results of training the ASR commercial system on two English data sets, enriched by error analysis. The second part presents results of machine translation performed by online tool Google Translate for English and Croatian and Croatian-English language pairs. Human evaluation in terms of usability is conducted and internal consistency calculated by Cronbach's alpha coefficient, enriched by error analysis. Automatic evaluation is performed by WER (Word Error Rate) and PER (Position-independent word Error Rate) metrics, followed by investigation of Pearson's correlation with human evaluation.

Keywords : automatic machine translation, integrated language technologies, quality evaluation, speech recognition

Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development

Conference Location : Chicago, United States

Conference Dates : December 12-13, 2020