How Is a Machine-Translated Literary Text Organized in Coherence? An Analysis Based upon Theme-Rheme Structure

Authors : Jiang Niu, Yue Jiang

Abstract : With the ultimate goal to automatically generate translated texts with high quality, machine translation has made tremendous improvements. However, its translations of literary works are still plagued with problems in coherence, esp. the translation between distant language pairs. One of the causes of the problems is probably the lack of linguistic knowledge to be incorporated into the training of machine translation systems. In order to enable readers to better understand the problems of machine translation in coherence, to seek out the potential knowledge to be incorporated, and thus to improve the quality of machine translation products, this study applies Theme-Rheme structure to examine how a machine-translated literary text is organized and developed in terms of coherence. Theme-Rheme structure in Systemic Functional Linguistics is a useful tool for analysis of textual coherence. Theme is the departure point of a clause and Rheme is the rest of the clause. In a text, as Themes and Rhemes may be connected with each other in meaning, they form thematic and rhematic progressions throughout the text. Based on this structure, we can look into how a text is organized and developed in terms of coherence. Methodologically, we chose Chinese and English as the language pair to be studied. Specifically, we built a comparable corpus with two modes of English translations, viz. machine translation (MT) and human translation (HT) of one Chinese literary source text. The translated texts were annotated with Themes, Rhemes and their progressions throughout the texts. The annotated texts were analyzed from two respects, the different types of Themes functioning differently in achieving coherence, and the different types of thematic and rhematic progressions functioning differently in constructing texts. By analyzing and contrasting the two modes of translations, it is found that compared with the HT, 1) the MT features "pseudo-coherence", with lots of ill-connected fragments of information using "and"; 2) the MT system produces a static and less interconnected text that reads like a list; these two points, in turn, lead to the less coherent organization and development of the MT than that of the HT; 3) novel to traditional and previous studies, Rhemes do contribute to textual connection and coherence though less than Themes do and thus are worthy of notice in further studies. Hence, the findings suggest that Theme-Rheme structure be applied to measuring and assessing the coherence of machine translation, to being incorporated into the training of the machine translation system, and Rheme be taken into account when studying the textual coherence of both MT and HT.
Keywords : coherence, corpus-based, literary translation, machine translation, Theme-Rheme structure
Conference Title : ICT 2021 : International Conference on Translation
Conference Location : London, United Kingdom
Conference Dates : October 21-22, 2021