

以進階的文句推薦方式使用語料庫做為英文寫作之輔助

摘要

對於大部分的人而言，寫作是一種深度的表達過程，需要詳盡而深入的描述能力及精準而嚴謹的語意呈現。對於非英語為母語的學習者(ESL/EFL)來說，英語寫作尤其是一個困難的過程，常常會因為單字、搭配字(collocation)、句型結構等方面的認知不足，或是受到母語認知的牽制影響，而造成用詞、語法、甚至語意的錯誤。

近年來，語料庫的發展帶來了豐富的語言資源，不僅可以對語言的使用提供許多統計分析上的資訊，也可以做為語言學習者在特定詞語使用上的參照對象。目前語料庫的詞語使用參照以 concordance 技術為主，提供以特定字詞為基準的上下文例句排列，這種功能對 ESL/EFL 學習者於寫作過程所提供的參照仍然相當有限；而對於不同的查詢條件，若沒有良善的輸入和比對機制，往往搜尋上的數量和例句品質無法滿足學習者的參照需求；除此之外，對於檢索結果語料庫也沒有評估與排序的概念，學習者往往需要在大量的資訊中篩選出有用的資訊；綜觀以上幾點，目前語料庫技術無法滿足不同程度的 ESL/EFL 寫作者在參照協助上的需求。

本研究提出一種文句推薦方法，針對 ESL/EFL 寫作者在寫作過程上的認知不足或不確定，從語料庫中尋找出可用的參照例句，進而提供針對性的協助。我們的文句推薦方法包含三個模組，第一模組是一個彈性的表達元素模組，能針對寫作過程中的語言資訊需求，以規定的表達元素呈現作者的認知需求。第二模組是一個檢索模組，指從語料庫中比對尋找符合作者語意需求的例句。第三模組是一個排序的模組，針對找出的例句，評估其符合作者語意需求的程度，並將之排序，以提升例句參照的使用效率。

本研究依據文句推薦方法實做雛形系統，以 British National Corpus 和科學人雜誌語料庫(Science America)當例句來源，並依據論文的客觀評估和學習者的問卷調查兩種評估方式來評量雛形系統的成效。經由上述兩種評估方式，本研究驗

證雛形系統能針對 ESL/EFL 於寫作過程中給予一定程度的協助成效，證明本研究
的文句推薦方法確實能達成寫作協助的既定目標。

關鍵字：Writing Assistance、Corpora、Information Retrieval、Concordance、
Recommender System

A Sentence Recommendation Approach to Using Corpora for English Writing Assistance

Abstract

For most people, writing is a process of profundity. It requires well description of capabilities and excellent semantic of representations. For ESL/EFL (English as a Second Language/English as a Foreign Language) learners, English writing is a particularly difficult process. They have error in words, syntax or semantics because their expressions are constrained by their mother tongue or by lacking of the awareness of vocabularies, collocation, or sentence-structures.

In recent years, according to the development of corpus technique, we have rich resource in language. The resource can not only provide many statistical analyses of informations in language research but also can be used as a reference in the use of specific words for the learners. Current corpus technique use concordance which displays the words with their surrounding text to present sentences, but that approach provides very limited references in the writing process for ESL/EFL learners. Also considering the different inquired queries, if we have no well-defined input and retrieval module, the quantity and quality of results could not meet the demand for the learners. In addition, corpus technique doesn't provide the ranking and evaluating skills to the retrieval results. As a result of that, learners often require filtering a large number of information to find something useful. In view of the above points, current corpus technique is unable to satisfy different degrees of ESL/EFL learners in the writing process.

This paper presents a Sentence Recommendation Approach technique. With regard to the inadequate knowledge in the writing process for ESL/EFL learners, we search for useful sentences from the corpus and provide them to learners. Our Sentence Recommendation Approach technique has three modules: one is Expression Element Module. It allows the learners to use some expression elements to represent their demand. Another is Retrieval Module. It means the process of searching the corresponding sentences based on the expression elements in Expression Element Module. The last is Sort Module. It means the process of ranking the results derived from the Retrieval Module.

Our research establishes the experimental system to verify the performance of our Sentence Recommendation Approach technique. We use British National Corpus and Science America Corpus for the source of sentences. The evaluation is divided

into our objective evaluation and the questionnaire evaluation. Both of them prove that our experimental system does some favor in the writing process for ESL/EFL learners. In other words, our Sentence Recommendation Approach technique really helps the learners in the writing process.

keyword: Writing Assistance 、 Corpora 、 Information Retrieval 、 Concordance 、
Recommender System