Field/分野：ESP

Using an Ad Hoc Corpus to Create an ESP Vocabulary List for Students Majoring in Tourism

Fujita, Reiko (Tokyo Keizai University)

The Japanese government’s emphasis on the tourism sector of the economy is reflected in its recent establishment of the Japan Tourism Agency (Kanko-Cho), within the Ministry of Land, Industry, Transport and Tourism. As the number of tourists increases, the demand for staff with tourism credentials will also rise. Noticing this trend, universities have already begun expanding their tourism studies programs.

This presents a dilemma for those charged with teaching English to these students. Since research on English for Tourism (EFT) is still in the developmental stage, there is little empirical foundation for designing ESP materials tailored to tourism majors. In particular, specialized vocabulary lists are one of the most important resources for ESP classes. Specialized ESP word lists have already been produced in fields such as Business English, Nursing, Medicine, and Engineering using field-specific corpuses. However, since the presenter has not been able to find any tourism-specific corpuses nor empirically based EFT word lists, she decided to create her own word lists using ad hoc corpuses generated from web pages.

Previous needs analyses (Fujita 2004, 2008) indicate that tourism majors have several different kinds of needs for English. First, students are required to read articles or books in English as part of their studies. Second, tourism majors are typically interested in careers at airlines, travel companies and hotels, where they will have to use English for their jobs. Finally, they also need to read journals and papers written in English for their own research.

In order to produce empirically-based, field-specific word lists for EFT, the presenter created four corpuses of written English in the following fields:
9月5日（土）14:40-15:40 ESP 9 D31

1. Air travel
2. Accommodations
3. Travel related firms
4. Tourism related academic articles

(approximately 10,000 words for each topic)

Based on her experience and expertise, the presenter chose the texts from web pages that seemed representative of the field. For example, web pages of several airlines and airports were selected to create a corpus for the field of air travel.

One corpus was created for each topic, and each was sorted using the AntConc 3.2 (www.antlab.sci.waseda.ac.jp/antconc_index.html) concordance software to create a word list. Then, in order to extract field-specific keywords from each word list, the Log Likelihood Ratio (LLR) was applied to compare the word lists to a General English corpus in order to eliminate general vocabulary, and to identify intermediate level vocabulary suitable for university students (Chujo 2006). The result was four lists of key words: one for each field.

These keywords lists were carefully examined, and the remaining general vocabulary was eliminated manually. The resulting lists contained the 120 most frequently-used specialized words for each field. For high frequency words that appeared at least 15 times in a corpus, common collocations and clusters were also annotated. Finally, the four lists were combined to produce a more general list of the 150 most frequent tourism-related words.

Copies of these five wordlists and further details of the process of creating them will be provided during the presentation.

References: