

Towards the Development of the Latest Version of *Parallel Link*, a Tool for Searching Japanese-English and English-Japanese Parallel Corpora: Its Interface, Corpora Installed, Search Functions and Research Applications

Yasunori Nishina

Kobe Gakuin University and Doshisha University, Japan

This presentation introduces a word profiler dedicated to translation corpora of Japanese-English/English-Japanese, which the presenter has been engaged in for the past six years and is currently developing. The final aim of the development of this word profiler is to extract reliable translation units that can be utilised when compiling dictionaries and building translation databases, which have been subjectively worked on up to now. To achieve this goal, a number of existing parallel Japanese-English and English-Japanese corpora (13 kinds for Ver.2.00) were reconstructed from scratch and combined uniformly. A word profiler interface was then developed to make them searchable online.

Specifically, firstly, for the translation corpus, we decided to work with nine different Japanese-English and English-Japanese translation corpora (Ver.1.0) and 13 kinds (Ver.2.0) that are already available to the public free of charge. As these are all in separate formats, we unified their formats from scratch and prepared them so that all translation corpora can be skewered and searched.

Next, an improved version of the LWP (Lago Word Profiler) was developed and adopted as an online search tool to batch search this combined corpus of translations. It is capable of searching for patterns in headwords, extracting collocations based on various statistics and displaying example sentences for each sub-corpus. This time, the version 2.00 interface, which was developed over six years and released in last summer in 2024, is presented. The current version 2.0 provides searching patterns and collocations of Japanese content words (nouns, verbs, adjectives, adverbs, and pre-noun adjectival) and presenting those instances with their English translations. The new functions added in Ver. 2.0, such as the context display function, the distribution graph function for each corpus of collocations, the voice function, and the Google Translate link function, will be introduced. Development up to version 4.0 is also planned over the next five years.

In the future, we plan to add various search functions for English collocations as well, and a parallel concordancer that enables searches from both Japanese and English. We detail these upcoming revision of this tool.

The theoretical background of the translation unit is largely based on Teubert (2001). Finally, we would like to introduce some examples of the extraction of bilingual linguistic patterns between Japanese and English using this tool. In particular, I will

take the translation of Japanese verbs with strong metaphorical connotations as an example, visualise the translation units quantitatively and touch on their genre characteristics.