

**Project title:** Impact of a Terminology Database with a Triangular Structure for Cross-Lingual Information Retrieval

My project is mainly related with the terminology work used in the context of language management in foreign research activities and international publicity. The main purpose is to combine the terminology work with Cross-Lingual Information Retrieval (CLIR) methods in order to remove the inaccuracies from Japanese translations of Danish originated Named Entities (NE).

Problems behind this project are described as follows:

If country-specific terms are translated in a plurality of ways, it potentially is a source of confusion and can even lead to serious misunderstandings among readers. Such situations can arise from the interrelation of the following reasons:

1) Due to the poor language resources between the Scandinavian languages and Japanese, translation works for international publicity and diplomatic documents often require the transitive translation using English as pivot. However, direct translations also exist especially when messages are sent directly from local translators residing in the Scandinavian countries. To illustrate this, I interviewed the Japan External Trade Organization (JETRO), and the Danish and the Japanese embassies on how to translate “Fødevarestyrelsen”. The answer from JETRO was “食品管理局 (The Food Administration)” directly translated from Danish, while the embassies answered “獣医食品管理局” through their official English translation of “The Veterinary and Food Administration”.

2) The Japanese phonetic symbol system (Katakana) cannot express precise pronunciations, and a Katakana loanword can thus have many expressions [Kubota 1993], e.g. “Göteborg” as “ヨーテボリ<sup>1)</sup>”, “イエーテボリ<sup>2)</sup>”, or “イエテボリ<sup>3)</sup>”, not derived from “Gothenburg” as in English. In contrast, Copenhagen<sup>4)</sup> as in English is commonly recognized as “コペンハーゲン”, not “クベンハウン (København)” as in Danish. This type of mixture of origins and pronunciations from a local language or from English is a matter of common practice in the Japanese language.

To solve this problem, I proposed in my Master thesis [Kano 2008] to develop a centralized NE terminology database with a triangular structure among Danish, Japanese and English as pivot (henceforth TDB). During the course of my PhD project, I would like to extract NE from various corpora, analyze term inconsistencies of Japanese translations obtained via direct translations and transitive translations using English as pivot, develop the TDB, and evaluate the results obtained from CLIR before and after the implementation of the TDB.