ENGLISH - THAI MACHINE TRANSLATION

by

Dr. Nitaya Kanchanawan
Department of Thai and Oriental Languages
Ramkhamkaeng University
Bangkok, Thailand

Paper presented at
the International Conference
on Thai Studies
22 - 24 August 1984
Bangkok
I Introduction

Studies and researches on Machine Translation (M.T.) or Automatised Translation (A.T.) have been conducted since 1949 in the United States and the USSR to translate English into Russian and vice versa. Later on M.T. of other pairs of languages have been carried on both in the USA and Europe. At present many countries in Asia, including Japan, China, Hong Kong, and Malaysia are working on M.T. as well.

Group d'Etudes pour la Traduction Automatique (G.E.T.A.) is the M.T. research group at Universite Scientifique et Medicale de Grenoble, France, under the leadership of Professor B. Vauquois, who has pioneered this undertaking since 1961. While M.T. projects in the USA have been delayed by the conclusion of a report from the Automatic Processing Advisory Committee (ALPAC), which consequently led American researchers to other related topics, e.g. Computational Linguistics, Artificial Intelligence (A.I.), and others, other research groups on M.T. in other countries, especially G.E.T.A., have never given up their effort.

Recently M.T. researches have been back to the main stream again. European Economic Community started to finance EUROTRA for an M.T. project. In Canada TAUM AVIATION project is undertaken at the University of Montreal, for instance.

G.E.T.A.'s translation system is called ARIANE System. The system has been developed for more than 10 years. Since 1978 the system has been tested to translate Russian into French. Other languages have since been tested, including Portuguese, Spanish, Italian, German, English, and Japanese. At present G.E.T.A. is expanding its project into industrial aspects, especially on English-French and French-English which is much in demand.

In 1980 Professor B. Vauquois introduced this system to Asean countries. The ARIANE System was installed at the Computer Center, Universiti Sains Malaysia, Penang, first. Then researchers in Thailand who are interested in this project have been contacted.

This committee, chaired by Professor Udom Warotamasikkhadit, is consisted of officials from three universities, namely Chulalongkorn University, Ramkhamhaeng University, and Prince of Songkla University (Haad Yai). Then on December 30, 1981 another order was issued to appoint a committee, chaired by Dr. Nitaya Kanchanawan, to undertake a research on Thai structures for English-Thai Machine Translation by the ARIANE System. This committee is supposed to work in cooperation with the first one. The new committee is composed of linguists and Thai language lecturers from six universities, namely Kasetsart University, Chulalongkorn University, Thammasat University, Ramkhamhaeng university, Silpakorn University, and Srinakharinwirot University.

This project aims at translations of science and technological texts from English into Thai by using The ARIANE System, starting from the results of morphological and structural analysis of English done by G.E.T.A.
II. The ARIANE System

The present version of the ARIANE System, ARIANE-78, has been running from 1978. The algorithmic components of the system may be illustrated as the following:

TRANSFER PHASE

LEXICAL TRANSFER

PRECEIVING TREE WITH TARGET LEXICAL UNITS

STRUCTURAL TRANSFER

RESULT ANALYSIS INTERMEDIATE STRUCTURE LABELLED TREE

MULTI-LEVEL ANALYSIS

INTERMEDIATE STRUCTURE FOR THE TARGET LANGUAGE LABELLED TREE

SYNTACTIC GENERATION

SURFACE REPRESENTATION OF THE TARGET SENTENCE: LABELLED TREE

MORPHOLOGICAL GENERATION

OUTPUT TEXT: STRING OF CHARACTERS