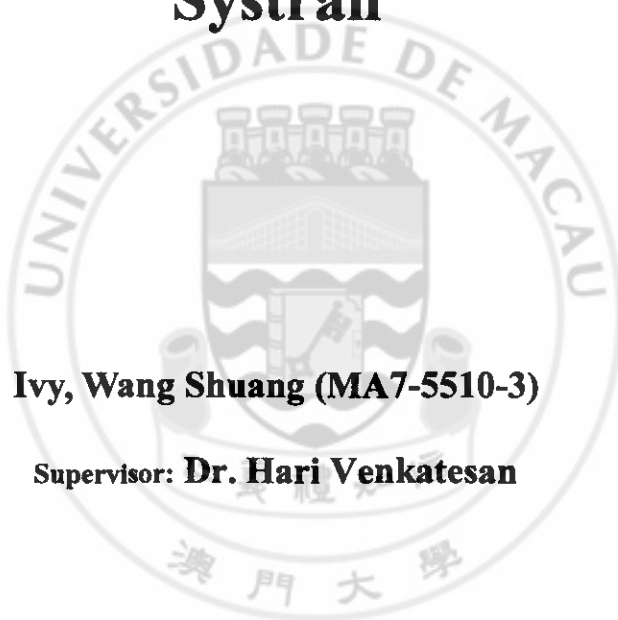


MFSH(APE)  
018  
WANG

# **A Critical Evaluation of Two On-Line Machine Translation Systems: Google & Systran**



**Ivy, Wang Shuang (MA7-5510-3)**

**Supervisor: Dr. Hari Venkatesan**

**Submitted to the Department of English  
Faculty of Social Sciences and Humanities  
in partial fulfilment of the requirements for the degree of  
Master of Arts  
at  
the University of Macau  
June, 2010**

## Contents

<b><u>Declaration.....</u></b>	<b><u>1</u></b>
<b><u>Acknowledgement.....</u></b>	<b><u>2</u></b>
<b><u>Abstract.....</u></b>	<b><u>4</u></b>
<b><u>Section 1 Introduction .....</u></b>	<b><u>5</u></b>
<b><u>Section 2 Literature review.....</u></b>	<b><u>10</u></b>
2.1 Three broad evaluation approaches.....	11
2.2 Studies related to declarative evaluation.....	12
2.3 Studies related to typological evaluation .....	14
2.4 Research issues .....	15
<b><u>Section 3 Methodology.....</u></b>	<b><u>18</u></b>
3.1 Parameters of evaluation.....	18
3.2 Research questions and test design.....	20
3.3 Methodological procedures.....	23
3.3.1 Test A1 .....	23
3.3.2 Test A2.....	24
3.3.3 Test B .....	26
<b><u>Section 4 Results and interpretation .....</u></b>	<b><u>28</u></b>
<b><u>Section 5 Conclusion.....</u></b>	<b><u>46</u></b>
<b><u>References.....</u></b>	<b><u>51</u></b>
<b><u>Appendix A: The three input texts for Google and Systran .....</u></b>	<b><u>52</u></b>
<b><u>Appendix B: Systran’s and Google’s output for the three input texts.....</u></b>	<b><u>54</u></b>
<b><u>Appendix C: Lists of testing points specific to each input text.....</u></b>	<b><u>57</u></b>
<b><u>Appendix D: Details of the guideline for evaluating English-Chinese machine translation.....</u></b>	<b><u>61</u></b>

## Abstract

Machine translation is by no means a new term today. And evaluation of machine translation, with a history of more than 50 years, has today developed into a project of great complexity. This is partly because of the built-in complexity of machine translation itself and partly because of the different purposes, goals and interests in different evaluations. This research aims to carry out a critical evaluation of two current on-line machine translation systems, Google and Systran. It is conducted in view of the fact that Google in 2007 developed its own machine translation system (statistics-based machine translation). Test material is used in order to make a comparison between the raw translations produced by the two systems. The analysis focuses on overall translation quality of each system and the type and number of errors arising in their output. The data show that Google shows relatively better performance than Systran as far as the tests in this research are concerned. On the other hand, the two systems' respective performance reveals that they each have different problems with respect to various linguistic elements.

**Key words:** MT Evaluation, Machine Translation, human rating, error analysis