Chapter 5
A Computational Cognitive Model of Human Translation Processes

Michael Carl
Copenhagen Business School, Denmark

ABSTRACT

Human translation process research analyzes the translation behavior of translators, such as memory and search strategies to solve translation problems, types of units that translators focus on, etc., identifies the temporal (and/or contextual) structure of those activities, and describes inter- and intra-personal variation. Various models have been developed that explain translators' behavior in terms of controlled and uncontrolled workspaces and with micro- and macro-translation strategies. However, only a few attempts have been made to ground and quantify translation process models in empirical user activity data. In order to close this gap, this chapter outlines a computational framework for a cognitive model of human translation. The authors investigate the structure of the translators' keystrokes and gaze data, discuss possibilities for their classification and visualization, and explain how a translation model can be grounded and trained on the empirical data. The insight gained from such a computational translation model not only enlarges our knowledge about human translation processes, but also has the potential to enhance the design of interactive MT systems and help interpret user activity data in human-MT system interaction.

1. INTRODUCTION

In recent years, MT has become widely available, covering many language pairs.1 Development for new language pairs is being increasingly short (Lewis, 2010) and the quality of the translation product increases based on the available resources and the similarity of the source and target languages.

However, to obtain high quality translations e.g., for dissemination, some kind of human intervention is necessary. In order to ensure the required quality of the translations and simultaneously increase translation production time,
numerous technologies exist or are experimentally implemented that ease human-machine interaction. A Machine Translation (MT) system may either work in a batch process as is the case in MT post- or pre-editing (e.g., controlled language translation) or in an interactive modus. In the case of interactive rule-based MT, the user interfaces of some MT systems (e.g., Systran, ProMT) allow the translator to extend or modify the lexical databases of the system at translation time; other systems interactively ask for disambiguation information (Boguslavsky, et al., 2005), which may be stored in a ‘companion’ file for later reuse (Choumane, et al., 2005). Recent implementations of interactive data-driven MT systems experiment with translation completion (TransType2, [Macklovitch, 2004]) or translation options (Koehn & Haddow, 2009).

Interactive rule-based MT systems ask for assistance to disambiguate the source text analysis by providing linguistic knowledge of the source language, whereas interactive data-driven MT systems ask the translator to disambiguate or chose from the generated translation proposals, thereby putting the user into the center of the translation process (Macklovitch, 2004). However, none of the approaches has yet led to a completely satisfactory solution to the issue of human computer interaction for MT.

Inadequate user interfaces, the translators’ feeling of competition and misrepresentation, and the lack of knowledge of their needs during translation have until now not led to a wider acceptance of the technology in the professional world. The challenge is not only how to improve machine translation technology itself so that it can adapt interactively to the translators’ needs, but also how to communicate, instrumentalize, and personalize the inherent value of MT for the stakeholder. It thus becomes important to study the human translator, who is at the center of the work and to understand their cognitive processes that take place during translation.

Current models of human translation processes (Göpferich, 2009; Hönig, 1991) aim at explaining and predicting translation behavior of novice and experienced translators, the types of units that translators focus on, memory and search strategies to solve translation problems, etc., and how these strategies are acquired. This work establishes and interprets in-depth empirically grounded knowledge of human translation processes to understand in detail how humans translate. Insights from empirical translation process research could then be carried over to the design of interactive MT systems, which could help interpret the user activity data in a way that would lead to improved human-machine interaction. Reactive and adaptable user interfaces can then be designed that can anticipate the translators’ needs to better support the translation processes.

In this chapter, we analyze and model the behavior of human translators when translating texts from scratch. A basic understanding of the unassisted translation processes is a prerequisite for developing more sophisticated systems. We have collected student and professional translator activity data—gaze and keyboard data—from 24 translation sessions. Each of the translations consists of three phases, which different translators realize to a different degree: an orientation phase, where a translator gets familiar with the text and activates subsets of the mental dictionary; a drafting phase where sequences of words are read and translated in a loop process; and a revision phase, where the translated text is re-read and revised.

This chapter focuses on formalizing and modeling the drafting phase: We fragment gaze and keyboard data into coherent units and analyze the amount of their overlap. Novice translators generally have a larger Source Text (ST) reading effort than experienced translators while experts show less fragmented typing behavior. Cognitive theories explain the reduced effort of expert performance by a higher degree of automatization of the mental processes, and the transformation of declarative knowledge into procedural knowl-
Related Content

Combinatory Categorial Grammar for Computer-Assisted Language Learning
[www.igi-global.com/chapter/combinatory-categorial-grammar-computer-assisted/64592?camid=4v1a](www.igi-global.com/chapter/combinatory-categorial-grammar-computer-assisted/64592?camid=4v1a)

Designing for Computational Expression: Four Principles for the Design of Learning Environments towards Computational Literacy
[www.igi-global.com/chapter/designing-for-computational-expression/108763?camid=4v1a](www.igi-global.com/chapter/designing-for-computational-expression/108763?camid=4v1a)

Computational Intelligence Using Type-2 Fuzzy Logic Framework
[www.igi-global.com/chapter/computational-intelligence-using-type-2-fuzzy-logic-framework/108722?camid=4v1a](www.igi-global.com/chapter/computational-intelligence-using-type-2-fuzzy-logic-framework/108722?camid=4v1a)

UNL-Based Bangla Machine Translation Framework
[www.igi-global.com/chapter/unl-based-bangla-machine-translation/78470?camid=4v1a](www.igi-global.com/chapter/unl-based-bangla-machine-translation/78470?camid=4v1a)