BENJAMINS

TRANSLATION LIBRARY

Methods and Strategies of Process Research

edited by
Cecilia Alvstad
Adelina Hild
Elisabet Tiselius

Methods and Strategies of Process Research

Benjamins Translation Library (BTL)

The BTL aims to stimulate research and training in translation and interpreting studies. The Library provides a forum for a variety of approaches (which may sometimes be conflicting) in a socio-cultural, historical, theoretical, applied and pedagogical context. The Library includes scholarly works, reference books, post-graduate text books and readers in the English language.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/btl

EST Subseries

The European Society for Translation Studies (EST) Subseries is a publication channel within the Library to optimize EST's function as a forum for the translation and interpreting research community. It promotes new trends in research, gives more visibility to young scholars' work, publicizes new research methods, makes available documents from EST, and reissues classical works in translation studies which do not exist in English or which are now out of print.

General Editor

Yves Gambier
University of Turku

Associate Editor

Miriam Shlesinger
Bar-Ilan University Israel

Honorary Editor

Gideon Toury
Tel Aviv University

Advisory Board

Rosemary Arrojo
Binghamton University

Michael Cronin
Dublin City University

Dirk Delabastita
FUNDP (University of Namur)

Daniel Gile
Université Paris 3 - Sorbonne Nouvelle

Amparo Hurtado Albir
Universitat Autònoma de Barcelona

Zuzana Jettmarová
Charles University of Prague

Alet Kruger
UNISA, South Africa

John Milton
University of São Paulo

Franz Pöchhacker
University of Vienna

Anthony Pym
Universitat Rovira i Virgili

Rosa Rabadán
University of León

Sherry Simon
Concordia University

Şehnaz Tahir Gürçaglar
Bogaziçi University

Maria Tymoczko
University of Massachusetts Amherst

Lawrence Venuti
Temple University

Michaela Wolf
University of Graz

Volume 94

Methods and Strategies of Process Research. Integrative approaches in Translation Studies
Edited by Cecilia Alvstad, Adelina Hild and Elisabet Tiselius

Methods and Strategies of Process Research

Integrative approaches in Translation Studies

Edited by

Cecilia Alvstad
University of Oslo

Adelina Hild
State University of New York

Elisabet Tiselius
University of Bergen/Stockholm University

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Methods and strategies of process research : integrative approaches in translation studies / edited by Cecilia Alvstad, Adelina Hild, Elisabet Tiselius.
p. cm. (Benjamins Translation Library, ISSN 0929-7316 ; v. 94)
"This volume of papers in honor of Birgitta Englund Dimitrova."
Includes bibliographical references and index.
1. Translating and interpreting--Research. 2. Psycholinguistics. I. Englund Dimitrova, Birgitta, 1946- II. Alvstad, Cecilia. III. Hild, Adelina. IV. Tiselius, Elisabet.
P306.2.M48 2011
418'.02--dc22 2011015157
ISBN 978 90 272 2442 2 (Hb ; alk. paper)
ISBN 978 90 272 8519 5 (Eb)

© 2011 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

17 papers inspired by the works of Birgitta Englund Dimitrova

Table of contents

Foreword **XI**

Methods and strategies of process research: Integrative approaches in Translation Studies **1**
Cecilia Alvstad, Adelina Hild and Elisabet Tiselius

Part I. Conceptual and methodological discussions

Interpreting in theory and practice: Reflections about an alleged gap **13**
Cecilia Wadensjö

Reflections on the literal translation hypothesis **23**
Andrew Chesterman

Tracking translators' keystrokes and eye movements with Translog **37**
Arnt Lykke Jakobsen

Seeing translation from inside the translator's mind **57**
Brian Mossop

Metonymic language use as a student translation problem: Towards a controlled psycholinguistic investigation **67**
Sonia Vandepitte and Robert J. Hartsuiker

Sight translation and speech disfluency: Performance analysis as a window to cognitive translation processes **93**
Gregory M. Shreve, Isabel Lacruz and Erik Angelone

Time lag in translation and interpreting: A methodological exploration **121**
Šárka Timarová, Barbara Dragsted and Inge Gorm Hansen

Part II. Process research in interpreting and translation

A new pair of glasses: Translation skills in secondary school 149
Miriam Shlesinger and Ruth Almog

Are primary conceptual metaphors easier to understand than complex conceptual metaphors? An investigation of the cognitive processes in metaphor comprehension 169
Antin Fougner Rydning and Christian Lachaud

Innovative subtitling: A reception study 187
Alexander Künzli and Maureen Ehrensberger-Dow

Errors, omissions and infelicities in broadcast interpreting: Preliminary findings from a case study 201
Daniel Gile

On cognitive processes during wordplay translation: Students translating adversarial humor 219
Ulf Norberg

"Can you ask her about chronic illnesses, diabetes and all that?" 231
Claudia Angelelli

Part III. Studies of interpreting and translation expertise

Effects of linguistic complexity on expert processing during simultaneous interpreting 249
Adelina Hild

Process and product in simultaneous interpreting: What they tell us about experience and expertise 269
Elisabet Tiselius and Gard B. Jenset

Developing professional thinking and acting within the field of interpreting 301
Gun-Viol Vik-Tuovinen

Results of the validation of the PACTE translation competence model: Translation problems and translation competence 317
PACTE Group

"This led me to start thinking about how this happened, and what the process behind it would be": An interview with Professor Birgitta Englund Dimitrova 345
Elisabet Tiselius

Publications by Birgitta Englund Dimitrova 361

Notes on contributors 367

Index 373

Foreword

Rarely is there either compensation or consolation for the missed opportunities that litter the path of a professional life. It was thus with great surprise and even greater pleasure that I received an invitation to write the foreword to this volume of papers in honor of Birgitta Englund Dimitrova. This is one of those rare compensatory moments.

Professor Englund Dimitrova's work has been central to the development of process-oriented research, and several of her publications have also addressed key methodological issues. Her willingness to cross traditional boundaries, such as those between product and process, between translation and interpretation, or between theory and practice, were early signs of a development that has since gained considerable momentum within the discipline. The breakdown of boundaries such as these is clearly evident in the papers collected here.

The volume of papers to follow is at the same time a testimony to the significance of Professor Englund Dimitrova's work as a translation process researcher and teacher and a snapshot of state-of-the-art methodological development and critique within this area of translation and interpreting research. The success in combining these two objectives is a considerable achievement on the part of the editors.

The studies presented here comprise a wealth of data types and methods, ranging from keystroke logs, eyetracking data, and time lag measurement, to questionnaires, interviews, tests, and retrospection, among others. From a broader perspective, methods include both quantitative and qualitative approaches and a variety of design types. In addition, mixed methods are used in several of the studies. There are a variety of interlingual tasks involved, ranging from written translation, to sight translation and interpreting. The wealth of empirical methods and data types brought to bear on process- and process/product-related issues is, in itself, a reassuring sign of the vitality of the field. The increasing methodological sophistication evidenced here, including the reflexive critique of own methods, is also an absolute necessity for the field.

On a more personal note, it is heartening to witness the generous participation of the numerous colleagues who have contributed in various capacities to the preparation of this volume. There is some deep justice in this: as a colleague, Birgitta has always been exceedingly generous with her own time and energy, often initiating projects and sharing ideas that are meant to create opportunities for others. Her generosity and enthusiasm are contagious: a number of her previous students and colleagues are now creating the same types of communities where they are. This too is a great gift to our research community.

It is with great respect for the achievements of Birgitta Englund Dimitrova, in whose honor this volume was prepared, and for the work of the contributing authors and the editors, that I commend this book to you, the readers.

Sandra L. Halverson
Bergen
February 2011

Methods and strategies of process research

Integrative approaches in Translation Studies

Cecilia Alvstad*, Adelina Hild# and Elisabet Tiselius† ‡
*University of Oslo, Norway / #SUNY Binghamton, USA /
†University of Bergen, Norway / ‡Stockholm University, Sweden

The objective of this volume is to present the reader with a diversity of innovative approaches to process research in Translation Studies. To this end, it addresses current theoretical and methodological concerns and also presents new empirical studies whose wide methodological range covers both translation and interpreting.

The central aim of process research is to "understand the nature of the cognitive processes involved in translating, with a focus on the individual translator" (Englund Dimitrova 2010a: 406). The process in translation and interpreting is generally divided into three phases: comprehension, transfer or shift between the two languages involved, and production (Englund Dimitrova 2010a). Naturally, none of the phases is directly observable, and researchers have therefore employed various methods for learning more about what goes on in translators' and interpreters' minds when they translate. To a certain extent it seems that process research has been methodology-driven rather than question-driven. Researchers have had to formulate their questions in relation to feasible methods. To a certain extent, methods have been imported from other disciplines, but they have also been added to and developed in relation to the researchers' concerns about what the relevant research issues are. This in turn creates a need for discussing methodological issues in relation to empirical studies.

Impelled by the difficulties of relying entirely on external observations and product analysis in studying the translation process, the pioneers of translation process research (such as Krings 1986; Tirkkonen-Condit 1989; Lörscher 1991; Jääskeläinen 1999; and Englund Dimitrova 2005) made use of think-aloud protocols, a method originally borrowed from cognitive psychology (Ericsson and Simon 1984). It is an introspective method in which subjects are expected to verbalize their mental processes while they perform a task. Norberg (in this volume) takes an innovative approach to this method, asking groups of students to translate together, and then studying their verbalized collaboration.

Soon afterwards another form of verbal reports – retrospection – was adapted by interpreting scholars because it allowed them to deal with the severe real-time constraints of the task. The subjects in experiments using retrospection are also

encouraged to verbalize what they do, but retrospectively, after the task is completed. In Interpreting Studies retrospection has been particularly popular in expertise research, and it has been used by Ivanova (1999, 2000), Vik-Tuovinen (2002) and Tiselius (in this volume). But it has also been used in studies on directionality (Bartłomiejczyk 2006; Chang and Schallert 2007) and on omissions in sign-language interpreting (Napier 2003; for retrospection in research on translation and the training of translators, see Hansen 2006).

In more recent years, new and more technologically advanced methodological tools for the study of the cognitive processes of translation and interpreting have been developed, e.g. eye-tracking devices and EEG. Process research has witnessed the emergence of mixed-method designs. Arnt Lykke Jakobsen's contribution to this volume describes a mixed-method design, namely, how the eye-tracker can be integrated with keystroke data. Alexander Künzli and Maureen Ehrensberger-Dow (also in this volume) combine eye-tracking and retrospective questions.

In spite of the improved methodological possibilities and more advanced theoretical frameworks for the study of the processes of translating and interpreting, only a modest number of experimental studies have been published on process-oriented research in translation and interpreting. Among these that do exist we would like to draw particular attention to the influential volumes *Tapping and Mapping the Processes of Translation and Interpreting* (edited by Tirkkonen-Condit and Jääskeläinen) in 2000 and *Triangulating Translation* (Alves 2003), as well as the important work presented recently in *Translation and Cognition* (Shreve and Angelone 2010), *Methodology, Technology and Innovation in Translation Process Research* (Göpferich, Alves and Mees 2010), without forgetting Englund Dimitrova's (2005) *Expertise and Explicitation in the Translation Process* quoted above. The results of the studies in these volumes play an important role in this legitimizing process.

New technologies make it more feasible to gain knowledge about the process today than thirty, twenty or even ten years ago, but the main reason for studying the process of translation and interpreting is of course not that it is possible but that it is important. Translators' and interpreters' behaviour can be assumed to be guided by both cognitive constraints (e.g. Halverson 2010 and other contributions to Shreve and Angelone 2010) and culturally acquired norms (Toury 1995). To know what parts of translation and interpreting are cognitively prompted and what parts are culturally acquired, we need to pay attention not only to the more traditional objects of study, i.e. translated texts and their cultural contexts, but also to the process, i.e. how they are brought about.

More knowledge about the process of translation and interpreting can have direct societal impact. With more knowledge about the process, about how individual translators and interpreters differ from each other (i.e. about translation and interpreting profiles), and with more knowledge about what the role of training and experience is for the translation and interpreting process, it is possible to develop more refined translation tools and training programs. This in turn may have also impact on quality and speed. The study of the translation and interpreting process can also contribute

more generally to the understanding of the human mind, since translation and interpreting are extremely demanding cognitive tasks that involve "comprehension in one language, production in another language, text production, use of aids, and problem-solving and decision-making" (Englund Dimitrova 2005: 6).

Kirsten Malmkjær has convincingly argued that once the field of Translation Studies has established its autonomy, "we cannot complete it without devoting a considerable amount of study to the central subject matter of Translation Studies, namely the processes of translation and the outcomes of those processes" (2000: 169). A wide range of aspects related to the process of translation remains to be explored. Furthermore, Sonja Tirkkonen-Condit advocated that "with the advance of process studies [...] it is necessary also to expose our methodologies to critical testing" (2002: 13). The present volume aims at taking one step further in filling these gaps.

The first part of this volume deals with a variety of methodological issues of process research in Translation and Interpreting Studies, e.g. integration of different data types in experimental studies of sight translation, evaluation of the reliability of eye-tracking and EVS measurements, and the desirability of combining ethnographic with cognitive research in interpreting.

The volume opens with a contribution by *Cecilia Wadensjö*. In a personal account, she starts with an overview of the development of Interpreting Studies, closely intertwined with her own career as a researcher. She examines trends in Interpreting Studies, which leads her to explore an issue very much present in the academic debate in Interpreting Studies today: Should Interpreting Studies try to bridge the gap between researchers and practitioners?

Andrew Chesterman engages in hypothesis-refining. He examines the literal translation hypothesis and assesses the importance the hypothesis may have against the background of two very different studies. The added value of the contribution resides not only in what it tells us about the literal translation hypothesis per se; it is additionally an excellent example of how one can build on the research conducted by others by framing it into a new and wider context.

Arnt Lykke Jakobsen gives an overview of the Translog project and how the eye-tracker was integrated with keystroke data and in turn led to the Eye-to-IT project. He starts with a thorough description of the design of the two projects and their relevance for each other. Through a case study he exemplifies the different challenges in interpreting the data elicited with these two methods. From the analysis of the keystroke and eye-tracking data he suggests the hypothesis of a recurrent processing pattern involving six steps which includes reading, typing and monitoring.

Brian Mossop suggests, from the point of view of the translator, that during translation one shifts between different ways of producing translational language, e.g. that a translator quotes, fudges, guesses or explicitates. He proposes that, while they are translating, translators tag their translations with predetermined vocabulary items such as GUESSING or FUDGING, thus signalling what their intentions are by engaging in self-analysis. Mossop additionally suggests that this kind of tagging may be of much use in a training situation, as the instructor would then know what the student's

intentions with a certain solution were. It may also be used in research as a complement to think-aloud protocols and retrospection.

Sonia Vandepitte and *Robert J. Hartsuiker* present the method of a novel kind of study: a controlled psycholinguistic experiment in which metonymic language use is tested as a student translation problem. Their point of departure is that there is a high emphasis on translation problems in Translation Studies research, and that linguistic competence modules are a standard component in translator training, in spite of the fact that we know very little about the extent to which crosslinguistic differences create process problems. The contribution explains the design of a test that showed how the translation of a particular metonymic language construction that differs across languages constitutes a translation process problem, both among untrained and trained students.

The most probable reason for speech disfluencies to arise in sight translation is the cognitive effort. This is what *Gregory Shreve, Isabelle Lacruz* and *Erik Angelone* report from their study on sight translation. Their purpose was to show that disfluency analysis, especially if triangulated with another form of data such as eye-tracking, can contribute to our interpretation of how the recognition of these difficulties and their resolution through repairs play out cognitively.

Early research on cognitive load has ushered in Translation Studies one of the most frequently used metrics – time lag – which continues to be refined and applied in studies of both translation and interpreting processes. *Šárka Timarová, Barbara Dragsted* and *Inge Gorm Hansen* turn their attention to the measure and offer a mix of meta-analysis and their own experimental findings addressing the reliability and validity of the measure for cognitive research. Their contribution has a very wide scope: it compares in a proximal manner methods used in previous translation and interpreting research to measure time lag, concluding that although they produced comparable results, certain differences in the application of the measure warrant further detailed experimental investigation. Their own research on eye-key span (EKS) in translation, using student-professional comparisons, shows a higher intra-subject variability, while suggesting potential for further studies of individual expertise-based differences. The researchers candidly discuss the problem of factoring out the effect of typing skill variations and highlight the need to fine-tune the tools in order to improve data quality and reliability. As regards time lag in interpreting, their investigation brings out the high inter- and intra-subject variability of the décalage pattern, which appears to be affected by a variety of external and internal factors (e.g. part of speech, position in the sentence).

The second part of this volume groups together reports on empirical translation and interpreting process investigations, with authors contributing articles on such topics as metaphor comprehension, between-group differences in the perception and resolution of translation problems by professionals and skilled L2 users, audience perception of different forms of subtitling, and the effects of teaching translation skills on L2 reading comprehension and meta-linguistic awareness.

Miriam Shlesinger and *Ruth Almog* are investigating the results of a two-year program in translation taught at high schools in Israel. English and Hebrew are used for teaching purposes but the focus lies on the translation process and the question: Do high school students who are taught translation develop their meta-linguistic awareness? The program has been running since the early nineties. The authors use questionnaires before and after the program as well as self-assessment and interviews. They conclude that meta-linguistic awareness among the students had developed considerably during the program.

Antin Fougner Rydning and *Christian Lachaud* investigate the cognitive processing of metaphor comprehension in a psycholinguistic experiment. They question whether primary metaphors are easier to translate than complex metaphors, referring to Lakoff and Johnson's (1999) hypothesis that primary conceptual metaphors are based on a simpler information processing architecture than complex conceptual metaphors. Fougner Rydning and Lachaud's results seem to corroborate Lakoff and Johnson's assertions by revealing greater conceptual clarity in the mind of their participants with primary metaphors than with complex metaphors. The relevance of their findings for translation research is further discussed.

In the subtitling industry and literature there is a commonly accepted norm, which says that a subtitle should only be one or two lines long and have a maximum of 80 characters. *Alexander Künzli* and *Maureen Ehrensberger-Dow* test the legitimacy of this norm in a study on innovative sub- and surtitling, which combined eye-tracking and retrospective on-screen questions intended to assess the reception capacity. According to the results, the presence of surtitles did not negatively affect the participants' reception capacity. In other words, the movie watchers were able to process a larger amount of textual information than is generally assumed.

On the assumption that simultaneous interpreters often work close to cognitive saturation (the Tightrope Hypothesis), which suggests that local phenomena are critical (Gile 2008), *Daniel Gile* analysed the output of interpreters who translated President Obama's inaugural speech live into French, German and Japanese by micro-units. The spread of errors and omissions (EOs), many of which affected simple speech segments, is in line with the Tightrope Hypothesis. The substantially higher number of EOs in the Japanese output is in line with the idea that interpreting difficulty is related to language-pair specific factors. The ratio of EOs to infelicities differed greatly among interpreters and suggests possible individual preferences for form or substance.

Ulf Norberg focuses on the importance of the source text genre for the quality of the translation process, a hitherto clearly under-researched area. He approaches this problem in a translation experiment conducted with translation students. Three small groups of students were assigned a translational task, to jointly translate a short text with a genre-typical translation difficulty. The translation difficulty was an adversarial wordplay without an evident counterpart in the language they translated into. The three different groups all discussed possible solutions, but none of the groups was able to produce a text as humoristic as the source text. The three groups employed several strategies, with one of these strategies, the one of paraphrasing the source text, clearly

leading to better results than other strategies. The discussion between the students showed that their problem with finding a target text solution they were satisfied with was not a problem of comprehension, or of recognizing the formal linguistic features which made the source text funny, but one of not being able to use this theoretical knowledge to produce a similarly humoristic text in their native language.

Claudia Angelelli takes on the process from a perspective of Critical Discourse Analysis. In a close reading of an interpreter-mediated nurse-patient encounter she shows how the interpreter exercises his power to alter the dynamics of the medical encounter. She also interviews the interpreter and finds that he considers it part of his job to make the patients understand and to obtain the answers to the questions put by the health care provider, even if it includes brokering questions or asking further questions on his own initiative.

The third part of this volume also consists of empirical studies and it is devoted exclusively to process-oriented research on the nature of translation and interpreting expertise.

Adelina Hild discusses different scientometric measures of linguistic complexity employed by psycholinguistic and educational research. In her study she lets two groups of interpreters (experts and novices) interpret two texts with different linguistic complexity. Her aim was to explore the effects of cognitive load, manipulated by increasing syntactic complexity, on simultaneous interpreting, and she concludes that expert processing is less vulnerable to linguistic complexity and that factors other than the syntactic properties of the ST are more important determinants of skilled SI performance.

In their study of interpreting expertise *Elisabet Tiselius* and *Gard Buen Jenset* sample interpreters at three different stages of their professional career: novices, professionals with short experience (two years), and professionals with long experience (average 25 years). The study combined a product analysis (evaluation of performance, using Carroll's scales of intelligibility and informativeness) with process protocols (retrospection, following Ivanova 1999). The findings corroborated previous research focussing on problems and strategy-use in translation, but also allowed the researchers to propose more fine-grained distinctions demonstrating the impact of experience (long vs. short) in the preference of strategies and perceived problems. By concretely outlining these differences and triangulating these findings with assessment of product quality, the study zooms in on new research areas which will allow us to begin to understand better the nature of expertise in interpreting.

Gun-Viol Vik-Tuovinen explores professional and tacit knowledge and how it is developed in the interpreting profession. Tacit knowledge, defined as the knowledge professionals develop through exercising their profession, is linked to interpreting strategies since both are necessary for interpreting. Strategies and tacit knowledge were investigated through questionnaires and formal interviews. The results stress situational aspects, knowledge and norms which are important entities in the process of professional development as an interpreter.

Research on translation expertise is also represented by new findings from the *PACTE Group*. Over the last decade the group has been carrying out empirical experimental research into translation competence in written translation and has developed a multi-vector model of translation competence. One of the six PACTE study variables, namely "identification and solution of translation problems", is presented here. Data analysis compares questionnaire response and interview protocols of 35 professional translators and 24 language teachers after performing direct and inverse translation in relation to four indicators: subject perception of the difficulty of the translation of the text, identification of translation problems, nature of the problems, and personal satisfaction with the solution. In relation to the first indicator, the results show a difference between teachers and translators: texts submitted for direct and inverse translation were perceived as subjectively more difficult by teachers than translators. The results related to the other indicators point to the importance of personality factors (confidence, self-critical attitude) on the subjects' perception of the nature and types of translation problems and their satisfaction with the solutions they produced in response to them. Findings lead the researchers to hypothesize that explaining translation problems requires a special kind of meta-knowledge.

All the abovementioned contributions have been written to honour Professor Birgitta Englund Dimitrova on her 65th birthday. An interview with Professor Englund Dimitrova follows after the scholarly contributions, and the closing chapter is a bibliography of her contribution to the field of translation and interpreting process research. Process research is an area which has been at the centre of interest for many Translation and Interpreting Studies scholars over the past twenty-five years, but particularly so for Birgitta. It is certainly suitable to honour her with a volume dedicated to an area where she has been very active and influential, especially since this area has not seen as many publications as could be expected. The authors are friends and colleagues of Birgitta and the volume therefore offers an interesting mix of contributions from influential translation and interpreting scholars, while also giving voice to the younger generation of Translation Studies researchers. Birgitta Englund Dimitrova herself points out the possible weaknesses of a volume of this kind: "[t]ribute volumes do not necessarily have any previously planned or imposed thematic structure" and furthermore, the fixed publication date "often creat[es] a tight deadline for both authors and editors", weaknesses that may result in "large variations in the size, scope and quality of the contributions" (2010b: 249). It has been our aim when preparing this volume to avoid these possible pitfalls and we hope that the volume will become what Birgitta Englund Dimitrova labels a "nursery of scholarship" (2010b: 257) with interesting, instructive and possibly pioneering works.

The work in this volume has given rise to many questions that we normally would have addressed to Birgitta herself. When consulting other colleagues instead, we have often received the answer, "Well, the person to ask about this would certainly be Birgitta." Her generosity, and the respect and warmth her colleagues feel for her has been of invaluable help in working on this volume. The mere mention of her name has been a veritable open sesame. Hence, in addition to the authors

represented in the volume, a large number of scholars have contributed to this volume and we would especially like to thank Ingrid Almqvist, Magdalena Bartłomiejczyk, Elisabeth Bladh, Chia-chien Chang, Agnieszka Chmiel, Helle Vrønning Dam, Ali Darwish, Jorge Díaz-Cintas, Peter Flynn, Emilia Iglesias Fernández, Yves Gambier, Riitta Jääskeläinen, Bente Jacobsen, Zuzana Jettmarová, Francis Jones, Donald Kiraly, Geoffrey Koby, Barbara Köpke, Pekka Kujamäki, Sigmund Kvam, Ritva Leppihalme, Yvonne Lindqvist, Minhua Liu, Ian Mason, Holly Mikkelson, Claudia Monacelli, Jeremy Munday, Presentación Padilla Benítez, Outi Paloposki, Jan Pedersen, Jonatan Pettersson, Franz Pöchhacker, Macarena Pradas Macías, Alessandra Riccardi, Hanna Risku, Andreas Romeborn, Anne Schjoldager, Candace Séguinot, Sonja Tirkkonen-Condit and Jeroen Vandaele. We are also grateful to Isja Conen, Sherida Deeprose, Stig Oppedal and the Institute for Interpreting and Translation Studies at Stockholm University for their support in the editing process.

This volume is our way of saying thank you, Birgitta, for everything you have done for us and for Translation Studies.

References

Alves, F. (ed.). 2003. *Triangulating Translation: Perspectives in Process-oriented Research*. Amsterdam/Philadelphia: John Benjamins.

Bartłomiejczyk, M. 2006. "Strategies of simultaneous interpreting and directionality." *Interpreting: International Journal of Research and Practice in Interpreting* 8 (2): 149–174.

Chang, C. and Schallert, D. 2007. "The impact of directionality on Chinese/English simultaneous interpreting." *Interpreting: International Journal of Research and Practice in Interpreting* 9 (2): 137–176.

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process.* Amsterdam/Philadelphia: John Benjamins.

———. 2010a. "Translation Process." In *Handbook of Translation Studies*, *Volume 1*, Y. Gambier and L. Doorslaer (eds), 406–411. Amsterdam/Philadelphia: John Benjamins.

———. 2010b. "Review of *Interpreting Studies and Beyond. A Tribute to Miriam Shlesinger* (F. Pöchhacker, A. Lykke Jakobsen and I. Mees (eds), 2007. Copenhagen: Samfundslitteratur) and of *Efforts and Models in Interpretering and TranslationRresearch: A Tribute to Daniel Gile* (G. Hansen, A. Chesterman and H. Gerzymisch-Arbogast (eds), 2008, Amsterdam/Philadelphia: John Benjamins)". *Interpreting* 12 (2): 249–258.

Ericsson, K. A. and Simon, H. A. 1984. *Protocol Analysis. Verbal Reports as Data*. Cambridge, Mass.: MIT Press.

Gile, D. 2008. "Local cognitive load in simultaneous interpreting and its implications for empirical research." *Forum* 6 (2): 59–77.

Göpferich, S., Alves, F. and Mees, I. M. (eds). 2010. *New Approaches in Translation Process Research* [Copenhagen Studies in Language 39]. Copenhagen: Samfundslitteratur.

Halverson, S. 2010. "Cognitive translation studies." In *Translation and Cognition*, G. M. Shreve and E. Angelone (eds), 349–369. Amsterdam/Philadelphia: John Benjamins.

Hansen, G. 2006. "Retrospection methods in translator training and translation research." *The Journal of Specialised Translation* 5: 2–41.

Ivanova, A. 1999. "Discourse processing during simultaneous interpreting: An expertise approach". (Doctoral dissertation, University of Cambridge, UK.)

———. 2000. "The use of retrospection in research on simultaneous interpretation." In *Tapping and Mapping the Processes of Translation and Interpretation*, S. Tirkkonen-Condit and R. Jääskeläinen (eds), 27–52. Amsterdam/Philadelphia: John Benjamins.

Jääskeläinen, R. 1999. *Tapping the Process. An Explorative Study of the Cognitive and Affective Factors Involved in Translating*. Joensuu: Joensuun Yliopisto.

Krings, H. P. 1986. *Was in den Köpfen von Übersetzern vorgeht*. Tübingen: Narr.

Lakoff, G. and Johnson, M. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought.* New York: Basic Books.

Lörscher, W. 1991. *Translation Performance, Translation Process, and Translation Strategies. A Psycholinguistic Investigation*. Tübingen: Narr.

Malmkjær, K. 2000. "Multidisciplinarity in process research." In *Tapping and Mapping the Processes of Translation and Interpretation,* S. Tirkkonen-Condit and R. Jääskeläinen (eds), 163–170. Amsterdam/Philadelphia: John Benjamins.

Napier, J. 2003. "A sociolinguistic analysis of the occurrence and types of omissions produced by Australian sign language–English interpreters". In *From Topic Boundaries to Omission: New Research on Interpretation*, M. Metzger, S. Collins, V. Dively and R. Shaw (eds), 99–153. Washington, D.C.: Gallaudet University Press.

Shreve, G. M. and Angelone, E. 2010. *Translation and Cognition.* Amsterdam/Philadelphia: John Benjamins.

Tirkkonen-Condit, S. 1989. "Professional versus non-professional translation: A think-aloud protocol study." In *The Translation Process*, C. Séguinot (ed.), 73–85. Toronto: H. G. Publications.

———. 2002. "Process research: State of the art and where to go next." *Across Languages and Cultures* 3 (1): 5–19.

Tirkkonen-Condit, S. and Jääskeläinen, R. 2000. *Tapping and Mapping the Processes of Translation and Interpreting: Outlooks on Empirical Research*. Amsterdam/Philadelphia: John Benjamins.

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam/Philadelphia: John Benjamins.

Vik-Tuovinen, G.-V. 2002. "Retrospection as a method of studying the process of simultaneous interpreting." In *Interpreting in the 21st Century*, G. Garzone and M. Viezzi (eds), 63–71. Amsterdam/Philadelphia: John Benjamins.

PART I

Conceptual and methodological discussions

Interpreting in theory and practice

Reflections about an alleged gap

Cecilia Wadensjö
Stockholm University, Sweden

During the past decades Interpreting Studies has developed quite substantially. A number of initiatives have also been taken to summarize and synthesize research on interpreting of various kinds, performed in a range of settings, involving spoken as well as signed languages (e.g. Garzone and Viezzi 2002; Pöchhacker and Shlesinger 2002; Hertog and van der Veer 2006). The aim of this paper is less to report on the results presented in this growing body of literature, and more to discuss the development of what can be seen as two competing trends – an urge for evidence-based interpreting practice on the one hand, and, on the other hand, an interest in studies of interpreting as an academic field in its own right, i.e. studies designed to explore and explain how interpreting happens. In response to an urge for narrowing a supposed gap between theory and practice, this paper suggests keeping a critical distance.

Keywords: interpreting research, evidence-based interpreting practice, interpreting as interaction, professional judgment

1. Introduction

During the past decades, Interpreting Studies has developed quite substantially as a professional and academic field. The development is far from clear-cut, however, and at times I find it hard to grasp in which directions the field is actually moving. Perhaps not surprisingly, the development in question has evolved around certain recurrent themes. A large part of the literature tends to repeat some "eternal" questions, concerning both the appropriate role(s) of interpreters and translation accuracy, both issues responding to the urge of policy makers to establish valid standards. Studies of interpreting have frequently applied a normative perspective, either taking certain normative behaviour for granted and/or aiming at demonstrating that interpreters violate established standards. Much recent research applies a non-normative perspective however, inspired both by appeals, e.g. Michael Cronin's influential call for a

"cultural turn" in Interpreting Studies (2002), and by concrete examples of solid research based on various kinds of interpreting data. Yet, the accumulated literature covering basic research, exploring interpreting as communicative, situated, cultural and cognitive practice, is still quite limited.

No doubt, questions of interpreting standards and quality relate to the advancement of interpreting as a professional field. But what about the relation of these questions to the advancement of Interpreting Studies as an academic field? The present paper takes this question as a point of departure.

1.1 Pros and cons of being a "practisearcher"

Like many colleagues exploring interpreting, I have a background as a practising interpreter. Hence, I could be labelled a "practisearcher", in Daniel Gile's (2000) term. In the late 1970s, I took my first authorization (in Sweden a certification is valid for five years). Having my diploma, I worked for several years as a freelance interpreter between Russian and Swedish. In the late 1980s, I had the chance to conduct doctoral studies on a project on interpreter-mediated communication in immigration hearings and medical encounters (Wadensjö 1992). Knowing the field from the inside, I had a number of questions and a strong motivation. Also, I had a large network of colleagues and other stakeholders who were willing to share experiences and information.

Naturally, I started off with a certain understanding of what interpreting is all about. Willingly or unwillingly, the educational programme – even if I thought of it as more or less a-theoretical – had socialized me into a certain conventionalized thinking about interpreting, put simply, as consisting of more or less difficult words and phrases in two languages, and as constituting a unidirectional, de-personified transfer from one language to another, on the basis of a view of communication as a unidirectional process from "sender" to "receiver" (cf. Reddy's (1979) idea of the conduit metaphor). Consequently, my image of an interpreter was in essence that of a re-sender (in new linguistic forms) of ready-made messages.

When I began my PhD project, my major ambition was to shed light on a marginalized but rapidly growing phenomenon in society. I wanted to make people in and outside of academia realize the complexity of the interpreter's job in the everyday settings I had experienced. My ambition was to explore dialogue interpreting – i.e. interpreting in face-to-face interaction – in its own right. I also had an ambition to make candidate interpreters realize how they ought to perform their professional role. It took me some time to realize that these two ambitions do not always go well together.

I set off with what may be called an everyday theory of interpreting, founded in an everyday understanding of language, communication and mind. Only when I began to put question marks around seemingly self-evident points of departure, e.g. to focus on the constellation of interacting people, what I later termed "a communicative pas-de-trois" (1998: 12), instead of on the individual interpreter, I started to develop a genuinely different understanding. I got to see things in a different way when I

stopped thinking "the interpreter must" or "must not" and began to explore "how does interpreting happen, more precisely?".

The co-ordinating aspect of interpreting (Wadensjö 1992, 1998) – the interpreter as manager of the discourse flow, as Cynthia Roy (2000) puts it – has of course been there all the time. No doubt, experienced interpreters have always been at least intuitively aware of the fact that in professional practice their interactional competence is of fundamental importance. As far as I know however, more detailed explorations of this competence started only in the early 1990s. Birgitta Englund Dimitrova (1991, 1997) was among the first translation scholars who focused specifically on turntaking in interpreted discourse. Exploring in detail two video-recorded medical encounters, she concludes that "the interpreter is in the centre of the turntaking process" (1997: 161). Pointing to some implications of the turntaking pattern, Englund Dimitrova raises open questions concerning what should and/or should not be counted as an interpreter's responsibility. In other words, she identifies relevant links between research and practice. Before discussing this relation further, I will devote some space to distinctions made between various types of interpreting.

1.2 Dialogue interpreting, community interpreting and conference interpreting

In English, interpreting performed in health and mental health care settings, at the police and other legal institutions, and in social security and immigration centres, is often termed community interpreting. In the UK, the cover term is more often public service interpreting. Dialogue interpreting can refer to interpreter-mediated interaction occurring also in other social settings. Moreover, this term emphasizes the social organization of interpreted communicative activity deriving from the circumstance that interpreting takes place in face-to-face encounters. Dialogue interpreting can be further defined in contrast to conference interpreting. Comparing the latter two concepts, one could point to the following essential features of conference interpreting (focusing on interpreting as shared activity):

- Primary speakers perform pre-planned (often written) speeches,
- Interpreters on duty speak in one, seldom in two languages,
- Interpreters perform from a separate booth,
- An (often multi-lingual) audience can follow the original talk and/or interpreters' talk through headphones.

The corresponding features of dialogue interpreting would be:

- Primary speakers are involved in spontaneous, spoken interaction,
- Interpreters speak in two languages, rarely in only one language,
- Interpreters perform in face-to-face situations,
- An interpreter is often alone in speaking and understanding both languages spoken (cf. also Mason 2000: 216).

In the educational programme for interpreters that I attended in the late 1970s, the distinction between conference interpreting and dialogue interpreting was not highlighted in the way it is today. The basic distinction made (focusing on the individual interpreter) was between consecutive and simultaneous modes of interpretation. Moreover, conference interpreting was seen as a model for interpreting generally, and research on interpreting often concentrated on the issue of (conference) interpreters' cognitive abilities and constraints. This might explain why theories of interpreting that were introduced in the training I attended were almost exclusively based on cognitive linguistics.

The alternative model of the interpreter-mediated conversation as a communicative pas-de-trois springs rather from pragmatic theory, from the idea of interpreters as active parties in constellations of people involved in common interaction. Also, it is based on what is known about communicative interaction performed between individuals speaking the same language, i.e. that we already as infants learn to communicate by taking and allocating turns at talking, using combinations of gestures and other non-verbal and verbal means, including words. We learn to interpret and respond to what others indicate by gazes, gestures, voice quality, vocabulary and so forth. It is hardly controversial to suggest that theses abilities constitute the basis for human interactive competence. Consequently, an interactive approach is valid in studies of conference interpreting, no less than in studies of dialogue interpreting.

Different models – interpreting as the individual interpreter's processing of information, and interpreting as individuals' interaction in interpreted conversations – have been used to generate new knowledge about interpreters' performance and about the nature of interpreter-mediated interaction. But how does this knowledge apply in practice?

2. Can theory help interpreters?

During the latest decades, the image of interpreting – conference as well as dialogue – has become both detailed and nuanced, thanks to research into naturally occurring interpreting activities, explored in relation to various contexts (interactional, institutional, socio-cultural, etc.). Some researchers are inclined to describe and explore not just how interpreters perform and interact, but also how they ought to perform. For instance, in the book Community Interpreting, Sandra Hale writes that "[m]ore research is needed on what the role(s) of interpreter ought to be, on whether such a role changes according to situation, participants and context and on the reasons behind the prescribed roles" (Hale 2007: 127). Calling for prescriptive research, Hale follows both an old tradition in the study of translation and interpretation and a quite recent call for evidence-based practice in various academic fields. The term evidence-based practice may be relatively new, but behind it lies an old drive. It links to a deeply felt wish among many people involved in research, to bridge the gap between theory and practice.

To my mind, discussions on the rights and responsibilities of interpreters are of fundamental importance in educational programmes at all levels. But what is the role of these issues in research on interpreted interaction? How do questions of interpreting standards relate to the development of interpreting as an academic field?

Looking at Translation Studies more generally, it is easy to find examples of a normative tradition. An obvious one would be Peter Newmark's book *Approaches to Translation* from the early 1980s, where he argues that "[t]he central concern of translation theory is to determine an appropriate method of translation" (Newmark 1981: 141). This book contains a lot of useful advice for practising translators and is still used in translator training. But does it generate any new knowledge, or any new research questions that may engage researchers in fields other than Translation Studies?

An interesting contribution to this debate appeared in 2002, in the book *Can Theory Help Translators?* by Andrew Chesterman and Emma Wagner. As the subtitle – *A Dialogue Between the Ivory Tower and the Wordface* – suggests, the book is designed as an exchange between "theory" and "practice". Wagner, a practising translator and translation manager at the European Commission in Luxembourg, starts her first chapter by asking what translation theory is about and expresses her doubts concerning whether theory can help translators to become better or improve their feeling of professional self-esteem. Chesterman, professor of translation theory at the University of Helsinki, counters by saying that he has a lot of sympathy for the assumption behind that question, but puts forward another question: should translation theory even have this aim?

Following her first sceptical view of the usefulness of "theory", Wagner then compares translators with doctors, a professional group that she believes earns public respect as professionals thanks to their knowledge in how to treat diseases, e.g. malaria. This seems like an appealing comparison to start with, but can the task of translators and interpreters be compared with that of curing malaria? To my mind, a potentially important difference here lies in the fact that doctors can be quite convinced that all medical practitioners share an ultimate goal. A translator and an interpreter can never be quite as sure, since communication – written as well as spoken – by nature involves a certain degree of vagueness and ambiguity, more or less open and hidden agendas, and shared, disparate and sometimes competing goals and assumptions.

2.1 Interpreting as communication

In face-to-face encounters, some of the vagueness and ambiguity are consequences of the distribution of turns at talking. As Cynthia Roy (2000: 84–100) shows, while scrutinizing an interpreted encounter involving English and American sign language, when primary participants introduce more or less simultaneously competing communicative projects, the interpreter, in order to secure a joint focus and a shared trajectory of talk, needs to foreground one participant's contribution at the expense of another's.

This problem is further explored in a study based on a dialogue-interpreted talk show interview (Wadensjö 2008). In this article, the interpreter's efficiency as an interpreter is shown to be a result partly of his ability to anticipate grammatical and pragmatic features of turn composition, partly of his fluency in English and Russian, and partly also of the interviewee's and the show host's communicative behaviour. While addressing one another as conversational partners and interacting with the studio audience and the viewers, the primary participants, Mikhail Gorbachev and Clive Anderson, cast the interpreter variably as sharing and as not fully sharing their ongoing exchange. The interpreter's ambiguous status as participant in the exchange creates a certain wiggle room, which all participants, including the interpreter Pavel Palazchenko, utilized as a communicative resource. Adapting to the specific, triadic communicative format, Gorbachev is successful in driving his communicative projects (telling anecdotes, stating his view of Boris Yeltsin as president of Russia, etc.). It might be tempting to conclude that the interpreter identified with Gorbachev, as it were, and supported his communicative projects rather than those of the BBC journalist (who seemingly had difficulties at times to steer the interview). But that would remain speculation. However, a close analysis of how the conversation unfolded moment-by-moment clearly shows that the interpreter's behaviour answered directly to the simultaneously occurring communicative behaviour of both parties.

These observations do not necessarily translate directly into rules and standards for interpreters. Indirectly, of course, it is probably helpful to realize that individuals' participation and non-participation in a social encounter – whether as interpreter or non-interpreter – is something interactively accomplished. Also, it might be important for practising interpreters to know that a powerful and talkative person, assisted by a carefully listening interpreter with a good sense of timing, has a good chance of standing forth as a competent conversational partner, also when he is talking in a foreign – in the context at hand – language. A reticent person is probably not less dependent on an interpreter with similar qualifications, rather the opposite.

Does this mean that research into how interpreting happens can help practising interpreters? Yes certainly, some would argue, it can help raise awareness, and it can assist interpreters and those who work with them in making professional judgements. No, not at all, others may say, thinking that a theory of interpreting is worth nothing if it fails to provide clear guidance for interpreters concerning how to perform interpretation. The latter claim is sometimes combined with arguments for stronger ties between interpreting research and interpreting practice, for "narrowing the gap".

3. Narrowing the gap?

Basically, the idea of narrowing the gap between research and practice is a positive one, it would seem. Clearly, it would be an enormous waste of time and money if the results of research on interpreting never had any relevance for the practice of interpreters. But I agree with Englund Dimitrova, who, with reference to Translation

Studies generally, argues that "we must allow for different kinds of research, treating all kinds of questions, such that are relevant to the professional practice and/or to society at large, as being one important type, but also such that are of theoretical interest, without any immediate applications in sight" (Englund Dimitrova 2008: 5). Just like in other academic fields, the distinction between basic and applied research can be used. If studies exploring how interpreting happens (in various modes and various settings) can inform practice indirectly, studies of how clients and professionals (lawyers, journalists, conference delegates, nurses, psychologists, etc.) experience the assistance of interpreters, and how interpreters experience their working conditions, would in principle be more immediately applicable to practice, for instance, when it comes to motivating changes and establishing new standards.

In this connection it is relevant to look at the very notion of evidence-based practice. For quite some time, evidence-based practice has been called for in medicine, were the concept was first applied. It has later been introduced in a number of academic fields, for instance education. When politicians demand better and more efficient schools, educators tend to demand researchers for better, more efficient, evidence-based training methods. Researchers are expected to produce scientific knowledge that informs teachers about e.g. which educational methods work and which do not work in class, or which methods teachers should apply to achieve the best results in terms of pupils' learning about, for instance, mathematics. And researchers in turn debate whether these expectations are reasonable.

For example, De Vries (1990), a Dutch professor in education, questions whether research can in fact produce the sort of knowledge that is needed in practice. De Vries makes a distinction between two roles that he thinks research can play, namely a technical role and a cultural role. In his view, research can on the one hand produce technical and instrumental knowledge. On the other hand, he argues, research can be important for practitioners also because it provides them with what he calls cultural knowledge, that is, different interpretations and understandings of practice. The latter is what he calls the cultural role of research.

Biesta (2007), also a professor in education, refers to De Vries (1990) and draws the discussion further, questioning whether research can produce any technical knowledge. His conclusion is quite simply no – "education is not a process of 'push and pull' […]. The effects of teaching are not produced in a causal or technological manner" (Biesta 2007: 298). Teaching techniques are, of course, important, but he argues that one cannot simply count on a causal effect between educators' activities and the learning that takes place.

In line with Biesta's reasoning concerning evidence-based educational practice, one could question whether more research into interpreting practice should be geared at finding, and ultimately controlling all the factors that determine a certain (effective/accurate) interpreting. Also, one may wonder if it is at all possible to identify and control all the factors that would guarantee a certain, optimal outcome of interpreter-mediated encounters, whether in face-to-face or in conference situations. Just like

education, interpreting is a symbolic activity; just like educational practice, interpreting practice consists of situations that are each time, in some sense, new and unique.

3.1 Possible and relevant links between theory and practice

So where could links between research and practice be established and developed? One option that immediately comes to mind is shared conferences. In the field of interpreting, The Critical Link is a series of conferences that has been established as a unique platform, providing a meeting place for practitioners and researchers working with interpreting in public service, healthcare and legal settings. Of course, communication between researchers, interpreter educators, interpreting-service providers and interpreters can be arranged also in other settings. Knowledge could be spread and cross-fertilized if researchers and educators working in the field of interpreting could visit educational programmes of other professionals, and vice versa. Reasonably, since many professionals today work in cross-cultural and cross-linguistic settings, it should not be difficult to find shared points of interest.

Research on interpreting will always, I believe, to some degree also be research for those who are professionally active as interpreters or interpreter educators. The question is, however, what this implies for the agenda of Interpreting Studies. How should it be specified and by whom? Is it up to researchers to establish standards? As a practising interpreter I definitely prefer thinking of professional standards as something to be discussed primarily within the interpreting profession. Research, on the other hand, can assist in interpreter's professional judgment and action, and it can raise professionals' awareness and support a better understanding of the complexities and conditions pertaining to interpreted interaction.

As a practising interpreter and interpreter educator, I certainly think it is important to raise general awareness concerning the kind of competences interpreting requires. As a practising researcher in the field of interpreting I find it necessary to analyse and problematize requests or requirements voiced among interpreters, interpreter educators, service providers and others who have opinions about what interpreters must or must not do. Also, I view it as my duty to raise issues that might be of interest also outside of Interpreting Studies. As has been said before, and Birgitta recalled in a discussion over coffee: "A strong discipline is one that feeds into other disciplines" (Englund Dimitrova 2010).

References

Biesta, G. 2007. "Bridging the gap between educational research and educational practice: The need for critical distance." *Educational Research and Evaluation* Vol. 13, No. 3, 295–301.

Chesterman, A. and Wagner, E. 2002. *Can Theory Help Translators? A Dialogue between the Ivory Tower and the Wordface.* Manchester: St. Jerome Publishing.

Cronin, M. 2002. "The Empire talks back: Orality, heteronomy and the cultural turn in interpreting studies." In *The Interpreting Studies Reader,* F. Pöchhacker and M. Shlesinger (eds), 387–397. London/New York: Routledge.

De Vries, G. H. 1990. *De ontwikkeling van wetenschap* [The development of science]. Groningen, The Netherlands: RUL/RUG.

Englund Dimitrova, B. 1991. *När två samtalar genom en tredje. Interaktion och icke-verbal kommunikation i medicinska möten med tolk.* Rapporter om tvåspråkighet, No 7. [Eng: Reports on bilingualism, No. 7] Stockholm: Stockholms Universitet, Centrum för Tvåspråkighetsforskning.

———. 1997. "Degree of interpreter responsibility in the interaction process in community interpreting." In *The Critical Link: Interpreters in the Community*, S. E. R Carr, A. Roberts, Dufour and D. Steyn (eds), 147–164. Amsterdam/Philadelphia: John Benjamins.

———. 2008. "Who am I – and does it matter? On the researcher in the research process in Translation Studies." (Unpublished manuscript).

———. 2010. Personal communication.

Garzone, G. and Viezzi, M. (eds). 2002. *Interpreting in the 21st Century: Challenges and Opportunities.* Amsterdam/Philadelphia: John Benjamins.

Gile, D. 2000. "Issues in interdisciplinary research into conference interpreting." In *Language Processing and Simultaneous Interpreting. Interdisciplinary Perspectives,* B. Englund Dimitrova and K. Hyltenstam (eds), 89–106. Amsterdam/Philadelphia: John Benjamins.

Hale, S. B. 2007. *Community Interpreting.* Basingstoke: Palgrave Macmillan.

Hertog, E. and van der Veer, B. (eds). 2006. *Taking Stock: Research and Methodology in Community Interpreting*. Linguistica Antverpiensia, New Series, 5/2006. Antwerp: Hoger Instituut voor Vertalers en Tolken, Hogeschool Antwerpen.

Mason, I. 2000. "Models and methods in dialogue interpreting research." In *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*, M. Olohan (ed.), 215–231. Manchester: St. Jerome Publishing.

Newmark, P. 1981. *Approaches to Translation.* Oxford: Pergamon.

Pöchhacker, F. and Shlesinger, M. (eds). *The Interpreting Studies Reader*. London/New York: Routledge.

Reddy, M. 1979. "The conduit metaphor: A case of frame conflict in our language about language." In *Metaphor and Thought,* A. Ortony (ed.), 284–324. Cambridge: Cambridge University Press.

Roy, C. B. 2000. *Interpreting as a Discourse Process.* New York/Oxford: Oxford University Press.

Wadensjö, C. 1992. "Interpreting as interaction: On dialogue interpreting in immigration hearings and medical encounters." (Doctoral dissertation, [Linköping Studies in Arts and Science No. 83] Linköping University, Sweden.)

———. 1998. *Interpreting as Interaction*. London/New York: Longman.

———. 2008. "In and off the show: Co-constructing 'invisibility' in an interpreter-mediated talk show interview." *META, Journal des traducteurs* 53 (1): 184–203.

Reflections on the literal translation hypothesis

Andrew Chesterman
University of Helsinki, Finland

This paper examines the well-known literal translation hypothesis and discusses its significance for translation theory. The hypothesis claims that as translators process a given text chunk, they tend to start from a literal version of the target text, and then work towards a freer version. The idea has been implied or explicitly studied by many scholars, and does not seem to have a single source.

After some preliminary conceptual analysis an optimal formulation of the hypothesis is proposed. The paper then assesses the hypothesis in terms of the kinds of wider significance any hypothesis can have. The criteria discussed are testability, relations with other hypotheses, applicability, surprise value and explanatory power. Some of Englund Dimitrova's research (2005) on the hypothesis is discussed. A rather different study by Nordman (2009) is argued to have implications for the broader contextualization of the hypothesis.

Keywords: literal translation, hypothesis, revision, research methodology

1. Introduction

Research projects in Translation Studies often either start with a hypothesis which they set out to test, or aim to generate one based on argument and/or data analysis. Some of these hypotheses attract wider attention, and they are then subject to further testing, either in replication studies with similar data or in other kinds of conditions. These hypotheses are usually dignified by an easily citable name, with the definite article in English: "the X hypothesis". Insofar as such hypotheses are widely corroborated, and probably refined and conditioned as well, they may eventually become part of the fundamental structure of a general theory of translation.

In the past two or three decades many interesting hypotheses have been proposed and tested in connection with the search for translation universals. Examples are the explicitation hypothesis, the simplification hypothesis and the unique items hypothesis (see e.g. the articles collected in Mauranen and Kujamäki 2004). These hypotheses make claims about very general features of translations as textual products, and

obviously have implications concerning underlying cognitive process(es). But there are also a few hypotheses that make direct claims about the translation process itself, such as the literal translation hypothesis to be examined here.

I will first explore some of the conceptual background underlying this hypothesis, and propose an explicit formulation of it. I then proceed to examine the significance of the hypothesis in terms of a number of basic criteria. The argument will be that on most of these criteria, the literal translation hypothesis is indeed rather important. After this general assessment two recent studies are briefly reviewed, in order to show possible new lines of research.

2. Conceptual background

Like many other terms in Translation Studies, "literal translation" has meant different things to different people (see Shuttleworth and Cowie 1997: s.v.). There is its vague general sense in opposition to "free" translation. One interpretation more or less equates a literal translation with a word-for-word one that is minimally adjusted to meet the demands of grammaticality (e.g. Catford 1965: 25). For some, a literal translation is ungrammatical; for others, it is grammatical, although it may still sound odd: cf. Vinay and Darbelnet (1958), where literal translation is given as one procedure implementing a "direct" translation strategy, alongside the use of a loan or calque. Like Vinay and Darbelnet, many scholars list literal translation as a possible technique or procedure which a translator can use when the conditions are appropriate, either in general or at some specific point in a text. Indeed, for some (e.g. Newmark 1988: 68–69), literal translation is seen as a default solution, from which a translator deviates only when, for some reason or other, a literal translation is not appropriate. I will start by taking a literal translation as one that is formally close to its source but nevertheless grammatical.

The brief list of varying views above illustrates one point where agreement is easy. Applied to a translation solution, the term "literal translation" may refer either to a complete text or to a translated fragment of a text. Similarly, applied to the translator's decision-making, literal translation may be a global or a local strategy. A given translation may thus be generally free, but contain elements that are literal; and a generally literal translation may contain elements that are not literal. (Cf. Englund Dimitrova 2005: 51–52.)

Literal translation is also closely related to some other terms that are used both in translation research and in contrastive analysis: equivalence, correspondence, congruence. Catford (1965: 32) distinguishes between (textual, or situational) equivalence and formal correspondence: this latter is defined as holding between source-language and target-language "categories" if they "occupy, as nearly as possible, the 'same' place in the economy" of the two languages. (For Catford, following Halliday, the categories are those of unit, structure, class and system.) This concept is thus a relative one, in the sense that the sameness is not absolute.

Catford's conceptualization of formal correspondence is not without its problems. There are different ways of measuring the degree of formal closeness, for instance, and it is not clear whether one should have priority over another. And it should be noted that his concept is defined in terms of languages, i.e. language systems, not texts. We can nevertheless infer that the notion of formal correspondence could form the basis of a possible definition of literal translation, as follows: a literal translation is one that shows the maximum of formal correspondence (within the constraints of the target language grammar). For any given source-text sentence or string, there would thus only be one translation in a given target language that would be called "literal": the one with the highest degree of formal correspondence. This restriction, however, is rather a drawback. We would surely like to be able to say that, of a whole range of possible or actual translations of a given source-language string, several are literal, and some are more literal than others, although version X is the *most* literal (with degree of literalness measured in some specified way).

The field of contrastive analysis also makes use of notions of equivalence and correspondence, and sometimes also of congruence. Various kinds of equivalence are recognized (see e.g. Krzeszowski 1990: 23f.), including translation equivalence. "Congruence" is a term used by some contrastivists, particularly those working within a generative framework, to refer to a relation between grammatical sentences in two languages which are semantically equivalent and "consist of the same number of lexical words, representing equivalent grammatical categories, arranged in the same linear order" (Krzeszowski 1990: 135). This too is a highly restricted notion. If we imagine a source string and a set of possible target translations, congruence in this sense would not necessarily hold for any of them, although we might still wish to say that some were more literal than others.

Both formal correspondence and congruence define states of affairs towards one end of a continuum of *formal similarity*. Formal correspondence between individual categories seems to indicate an even higher degree of similarity than congruence between strings or sentences, in that a part of a sentence might be formally corresponding although the whole sentence is not congruent. Beyond these cases we can imagine even closer matches, but they would no longer be grammatical.

What about the other end of the continuum, the non-literal end? There are obviously translations that bear no formal similarity at all with the source (although they must surely bear some other kind of relevant similarity if they are to be called translations at all). But there are also many translations that are formally similar only, for instance, to the extent that the target string is, like the source string, a sentence. Such sentence-similar translations, however, may well be non-literal in most other respects. Or, at a higher level, what about translations that bear a formal similarity only at the level of paragraph, or section, or chapter, or even at the level of the complete text? The conclusion I draw is that, in order to be useful, the notion of literal translation must be a comparative one. We can thus say that in given respects, translation A is *more literal* (i.e. is formally more similar to the source, on given criteria) than translation B. True, criteria vary and may conflict, so they need to be defined in each case. However, there

seems no obvious borderline between literal and non-literal translations; they are not two distinct sets. This conclusion applies both to the notion as referring to the text as a whole and as referring to a local strategy. Our initial assumption must thus be refined a bit: a literal translation is one that is formally *closer* to its source than some other translation of the same source chunk, on given criteria. (See the example below.)

There is one further conceptual complication, having to do with the relation between different kinds of formal similarity on the one hand and equivalence on the other. For Catford, formal correspondence does not imply (textual) equivalence. For Krzeszowski, however, congruent strings are a subset of (translationally, or semantically) equivalent strings. In Translation Studies, we have long argued about the notion of equivalence, and I will not repeat the debate here. (For a recent survey, see Pym 2010.) In the present context it will suffice to follow Krzeszowski's general line: translation research looks at the creation of equivalence, the various forms it can take, and what effects these forms may have. "Equivalence" here can be glossed as "relevant similarity". On this view, it makes sense to restrict the kind of formal similarity we are interested in to cases that occur within the bounds of this translational equivalence. Data that may show some formal similarity but no evidence of any translational equivalence are not relevant to the present argument (although they might well be relevant to language typologists or contrastivists, of course).

3. The hypothesis and its significance

The *literal translation hypothesis* does not refer synchronically to the final product of the translation process, nor to local instances within that product. It makes a diachronic claim about the translation process. In view of the discussion above I will formulate it as follows: *during the translation process, translators tend to proceed from more literal versions to less literal ones*. The underlying assumption is that the translator's cognitive processes will tend to be influenced, initially, by formal features of the source text. Note that this formulation of the hypothesis states a change in a certain direction. It does not start by postulating "a literal translation" as a starting-point of this change. There are two reasons for this. One is the difficulty of specifying an operational borderline between the range of possible translations that are somehow formally similar, and the range of translations that are not, as discussed above. Without a clear cut-off point, it is simpler to focus on the direction of change rather than the initial point of change. The second reason is that normally we are looking at a single translation, not multiple translations of the same text; and in any case we are not normally taking account of potential "shadow" translations that could have been used but were not. (The term "shadow translation" is from Matthiessen 2001: 83.) This means that in such cases we have no other translation with which to compare translation A. But by formulating the hypothesis in terms of a direction, we provide ourselves with a point of comparison: the hypothesis claims that initial (or earlier) draft version A is formally closer to the source than the later version B. – An alternative statement

of the hypothesis makes this explicit: in effect, we are talking about a *deliteralization* hypothesis, i.e. a move from more literal to less literal.

Let us consider a simple example, from my local Eniro telephone directory, which gives information in Finnish, Swedish and English. I assume that Finnish is the source language.

(1) Finnish: Kotimaan kaukopuheluita voit soittaa usean eri verkon kautta.
[domestic long-distance calls you can call via several different networks]

(2) Swedish: Du kan ringa inrikes fjärrsamtal via flera olika nät.
[you can ring domestic long-distance calls via several different networks]

(3) English: Long-distance calls in Finland can be made over several different networks.

A shadow English translation could have been:

(3a) You can make domestic long-distance calls via several different networks.

This structural choice would have been formally more similar to the Swedish version, with the same order of main constituents as the Swedish. The published English version (3) explicates "domestic" to *in Finland*, and introduces a shift from active to passive verb, but preserves the thematic order of the Finnish, starting with the *calls*. If we now compare the two English versions, we can say that (3) is more similar to the Finnish source on the criterion of preserving the thematic order, but (3a) is more similar on the criterion of preserving the voice of the verb phrase. With respect to the shift in (3) to *in Finland*, the literal translation hypothesis would claim that the translator first thought of the unshifted alternative *domestic*, as in (3a), but then adopted the less literal version in (3). Given explicit criteria, judgements of relative similarity are thus not difficult. However, in order to assess whether (3) or (3a) is formally more similar *overall* to the Finnish source, we would need some justified way of prioritizing different criteria. If syntactic structure is held to take precedence over thematic structure, the literal translation hypothesis could claim that (3a) as a whole occurred to the translator's mind first, as being syntactically close to the source, and was then modified in several respects: the shift to *in Finland*, the voice change, and also the change from literal *via* to *over*.

The hypothesis itself is by no means a new idea, of course. Toury (1995: 191) already cites Ivir (1981: 58) on the idea that translators start from target versions that show formal correspondence, and then move on to freer versions when they need to in order to achieve a relevant equivalence. The rejection of the initial literal version is assumed to be made by some kind of cognitive monitor, and several scholars have proposed a "monitor model" to represent this (see e.g. Tirkkonen-Condit 2005).

Before considering some recent empirical evidence I would like to bring up a different question: Why might the literal translation hypothesis matter? Why might it be significant? I will suggest several reasons. My intention here is also to indicate in

what respects the hypothesis can serve as a good methodological model in translation research.

First, it can be tested empirically. If we operationalize the verb "tend" in the above formulation of the hypothesis (e.g. to "during the translation process, translators in at least x% of cases studied proceed from more literal versions to less literal ones, rather than in the opposite direction"), the hypothesis can also be falsified. It is therefore vulnerable, which is a merit. Degrees of formal similarity can be operationalized in terms of the number and type of shifts. The most complex model of shift analysis so far proposed is that by Leuven-Zwart (1989/1990), but simpler measures could also serve the purpose.

Second, it can be tested in several quite different ways. This is also a merit, partly because it makes the hypothesis more vulnerable (and possible multiple corroboration is correspondingly more meaningful, especially if different analyses can be triangulated), and partly because it indicates that the hypothesis may have relevance to different research frameworks and thus perhaps encapsulate a fairly general insight. The various ways in which the hypothesis can be tested include the following (most of them are used or referred to in Englund Dimitrova 2005):

- Think-aloud protocols. Do translators' verbalizations show movement away from more literal versions? (See e.g. Englund Dimitrova 2005.)
- Keystroke logging analysis, such as Translog data (e.g. Jakobsen 1999).
- Interim solutions analysis (the study of the revision process across a series of drafts). This was the context of Toury's reference to Ivir, cited above.
- The study of repairs in simultaneous interpreting (cf. work referred to in Tirkkonen-Condit 2005). Interpreters appear to use fewer repairs when there is more syntactic similarity between strings in the two languages, which suggests easier processing.
- The study of the time taken to translate different kinds of idioms, some of which have formally matching versions in the target language (and tend to be translated faster) and some of which do not. (Also discussed in Tirkkonen-Condit 2005).
- The study of interference in general. Interference is of course a sign of some (usually) unwanted similarity that has been carried over from the source text into the target version.
- The study of differences between novice and professional or expert translators. A plausible corollary to the hypothesis would suggest that professionals and experts proceed more quickly, and further, along the path away from an initial literal translation, or that they actually start their processing from a less literal point. (See e.g. Englund Dimitrova 2005.)
- The study of translation performed under conditions of unusual time stress. One might expect that when processing time is strictly limited, more recourse is taken to literal versions, but research on this has so far been rather inconclusive (see e.g. Jansen and Jakobsen 2000).

If such tests are replicated with different language pairs and translation directions, one result might well be that the strength of the literal translation tendency varies under different language conditions, and indeed according to other kinds of conditions as well. (Cf. the survey of interference conditions in Murphy 2003.)

Third, the hypothesis is openly a counter-hypothesis to a competing claim: the deverbalization hypothesis, proposed and indeed assumed (but not empirically tested) by the so-called Paris school of interpreting, a claim implying the separation of form and meaning during processing. True, the deverbalization hypothesis was originally proposed for interpreting, but it has also been taken to apply to translation (see e.g. Seleskovitch and Lederer 1984). Nida's well-known river-crossing model of translation (e.g. 1964), comprising the three stages of analysis, transfer and restructuring, appears explicitly to assume deverbalization, at least insofar as the formal structure of the source text is initially reformulated at what Nida calls a "near-kernel" level of abstraction. In Nida's model, however, the initial deverbalizing move away from the source surface structure is represented as taking place within the source language, not the target language. Evidence in favour of the literal translation hypothesis would thus suggest some initial transfer to the target language *without* analysis, which would go against the model. Yet there might also be evidence of a move towards freer renderings during the restructuring process, which does take place within the target language. That said, it is clear that Nida's model is not based on explicit empirical evidence, and was presumably intended to have pedagogical and prescriptive priorities, as indeed was the deverbalization idea. Nevertheless, the fact that the literal translation hypothesis stands in a dialectic relation with a competing claim gives it a sharp theoretical relevance. It also has the rhetorical advantage of enabling scholars to formulate their discussions about it as a confrontational debate, as Englund Dimitrova does in her explicitation study (2005).

Fourth, the literal hypothesis can also be linked supportively to other current hypotheses in ways which bring it greater explanatory power, because these extend beyond linguistic description towards underlying cognitive processing. Two examples are the unique items hypothesis (Tirkkonen-Condit, e.g. 2004; Chesterman 2007) and the gravitational pull hypothesis (Halverson, e.g. 2003, 2007). The former is more specific than the literal translation hypothesis, and the latter is more general. Halverson's idea of gravitational pull focuses on how target-language category prototypes and superordinate conceptual schemata tend to influence the translator's choices, leading to the over-representation of certain kinds of items. This cognitive pull, argues Halverson, explains such putative translation universals as simplification, standardization and generalization. But underlying this idea there is obviously the assumption that salient source-text features will also exert a pull (i.e. leading to interference of some kind). In her discussion of the unique items hypothesis, Halverson (e.g. 2003: 223) makes this point explicitly. In the absence of any conceptual overlap between source and target structures, it is only to be expected that target-language-unique forms will be under-represented. In other words, if there is a choice between a target structure that is formally similar (and hence cognitively

salient at the moment of target-item selection) and one that is not, the translator will tend to select the formally similar one and thus save processing time and effort. Halverson also makes the important point that similar effects have been observed in studies on second language acquisition. This implies that so-called translation universals may not be specific to translation, but have to do more generally with language use under particular constraints.

Fifth, the hypothesis also has interesting potential applications in the description and explanation of individual translator styles, and perhaps in the optimization of revision procedures (see e.g. Mossop 2000, 2007). There may be more than one tendency at work: some translators (perhaps under certain working conditions, or with certain language pairs or translation directions or text types, or with certain personality types, or whatever) may tend to process in a *deliteralizing* direction, from more literal towards less literal, while others work in the opposite direction, beginning with a freer version and then pulling it back closer to the source text during processing or revision (i.e. *literalizing*).

In one respect, the literal translation hypothesis does not seem so significant. Because it appears to be highly plausible, it is a rather cautious hypothesis, not a bold one. Research results that went against it would really surprise us. Bolder, and potentially more interesting, sub-hypotheses might eventually emerge when we know more about the specific conditions under which a processing move from more to less literal tends to occur, and when it tends not to occur.

4. One recent study, and one different one

So what kind of empirical evidence has been found for or against the literal translation hypothesis? Englund Dimitrova (2005) has much of relevance to say about our hypothesis in her rich study on explicitation and translator expertise. Her subjects were three small groups of language and translation students and translation professionals, working from Russian into Swedish. The text to be translated was a short one, six paragraphs from an art album on the life and work of a Ukrainian poet and artist; a total of 30 sentences. And note that this was an experimental situation, not an authentic one.

Several research methods were used: TAP, videotaped keystroke logging, and revision analysis (interim solutions). This revision analysis covered both revisions made during drafting and those made during the post-writing stage. Englund Dimitrova discusses many of the points mentioned above, and her results certainly support the hypothesis (which is not actually stated as such, in this study), running clearly counter to the deverbalization hypothesis.

Englund Dimitrova's formulation of what she means by literal translation starts from the local sense of the term. For her, a literal translation is "a TT fragment which is structurally and semantically modelled upon the ST fragment while respecting TL grammatical constraints" (2005: 53). This formulation does not imply a single

"maximally" close version (i.e. *the* literal one), but allows, for a given fragment, a range of translations that could be said to be "modelled" on the source fragment, to different degrees and in different ways. It thus seems open to the borderline difficulty discussed above, and to the difficulty of a lacking point of comparison. But Englund Dimitrova is mainly interested here in explicitation, and her operationalization of a literal translation is one that is not explicitated (2005: 51). In this way, she implies one clear criterion for measuring the degree of formal similarity: an explicitated translation is less formally similar to the source than a non-explicitated one – and there is the point of comparison. In other words, although she starts by positing an initial literal translation, the operationalization of the decrease in formal similarity is done in terms of the presence or absence of explicitation shifts.

Her definition also includes some notion of semantic similarity ("structurally and semantically modelled"). One may debate the extent to which semantics is a matter of form, but her sense of "literal" does seem to be slightly wider than mine.

She found that there was "a tendency for syntactic revisions to result in structures that were more distant from the structure of the ST than the first version chosen" (2005: 121). This tendency was particularly evident in the revisions made by the group of professionals. The TAP data also showed that the subjects tended to be aware of the need to move away from literal versions. She wonders (2005: 146, 232) why professionals bother to write down a literal version at all, since it will probably be changed at a later stage. Perhaps this is a way of economizing on short-term memory. (*Pace* Nida, transfer can thus take place in the absence of any analysis, even in the case of professionals.) She also notes that literal versions are often the ones that are first verbalized, before anything is written down. This supports the idea that the literal version is a default solution. She acknowledges (2005: 148) that this conclusion is only based on an analysis of the revision of syntactic structures and does not take into account semantic or pragmatic changes. – However, if literalness is defined strictly as *formal* similarity, this limitation would be less problematic, although there would remain the challenge of sorting out priorities between potentially clashing formal criteria (as mentioned above).

In her conclusions, Englund Dimitrova underlines the importance, for professional translation competence, of knowing how best to exploit literal translation during the translation process. She found three prototypical patterns (2005: 234): (i) literal translations are formulated orally, then revised orally before a version is written; (ii) literal translations are initially written, then evaluated and revised during drafting; and (iii) literal translations are written as provisional solutions, and revised during post-drafting revision. She concludes that for professionals, the use of literal translation is thus a *processing* strategy. Its frequency no doubt varies according to the typological relation between source and target languages.

A brief aside is worth adding here. Machine translation, when not drawing on a huge translation memory of human translations, can also provide a literal translation, of a kind (cf. pattern (ii) above). Pym (2009) reports an exploratory study on how translator trainees can use a machine translation as a way of arriving at a fast first draft.

And here is a professional literary translator describing how he dealt with a translation of a family memoir from Italian to English: "Since Italian is not one of my working languages, I did the translation in two stages, the first of which was to run [the] Italian original through a commercial translation software called Systran" (Wilson 2009: 34). The result was a very literal version indeed, "not yet English, but very helpful", and he proceeded to polish it.

How valid are Englund Dimitrova's results? The study only used nine subjects, selected by convenience, and one short translation. As the author notes, it was only an initial, exploratory study. And there were other research questions which were more central to the project than the issue of literal translation. But with respect to the literal translation hypothesis, I find Englund Dimitrova's results rather convincing.

The second study I would like to mention briefly is very different, and does not explicitly test the literal translation hypothesis at all. However, I shall reinterpret one of its results and suggest that it shows how the hypothesis could be placed in a wider context, and also opens up a new perspective on it. The study in question is a PhD thesis on revision by Nordman (2009). Nordman's research examined in general how the revision system works for the Swedish translation of laws in the Finnish Parliament. (Laws are usually first drafted in Finnish and then translated into Swedish, Finland being a bilingual country.)

In Nordman's case, the revision in question is not self-revision but other-revision. The system she studied consists of a number of professional translators, revisers and legal experts, located in different offices (the Government translation unit, the Ministry of Justice Swedish revision unit, and the Swedish Department of the Finnish Parliament). Initial translations are passed up the chain, subject to different revisions at four different stages. So we are dealing with a case of institutional, team translation. The research project involved textual revision analysis, questionnaires and interviews. The texts in question are very different from those in Dimitrova Englund's study, and they have a different skopos. Here, we are dealing with legal texts, the drafts of bills that will become laws; and the revision situation is authentic. It is also relevant to point out that the translators studied were of course aware that their work would be revised, as normal. Despite all these differences between the two studies, I will argue that Nordman's work does suggest a new angle on the literal translation hypothesis.

Nordman analyses her revision data partly in terms of four norms: legal adequacy, correct translation, correct language and readability. Of these, "correct translation" has to do with formal and semantic similarity with the source. An example (2009: 158):

(1) Finnish source: (3) luoda edellytykset radiolaitteiden mahdollisimman vapaalle liikkuvuudelle […]
['to create conditions for the maximally free mobility of radio sets']

(2) Initial translation: (3) att skapa betingelser för att radioanläggningar ska få röra sig så fritt som möjligt […]
['… so that radio sets can be moved as freely as possible']

(3) Revised version: (3) att skapa betingelser för att en så fri rörlighet för radioanläggningar som möjligt […]
['… a mobility … that is as free as possible']

The other three norms relate to conformity to target language and genre, and to official recommendations for legal Swedish style, such as adding missing commas, or normalizing introductory structures in a list. At the four stages of revision, Nordman found that there was variation in the way corrections and changes could be ascribed to these different norms. At the first stage, the "correct translation" norm accounted for 15% of the revisions made. But – and this is the point I want to pick out – during the three later stages this proportion did not decrease; on the contrary, it rose to about 30% of the revisions made at each of these later stages (2009: 219).

This is surprising. One might have expected that adequate equivalence would have been established early in the revision process, with later changes having more to do with target language requirements (along the lines of Nida's restructuring, perhaps). That would have shown a deliteralizing process that continued through the revision phases. But what we seem to have here is evidence of *reliteralization*. Later revisers, in other words, have felt that the similarity relation is not close enough and altered the text accordingly. One reason for this may be the fact that in this context the source text is not entirely constant, being itself subject to some small changes as the text is finalized; but this seems to account for only some of the later equivalence revisions. Another reason may be the structural weakness of the whole revision system itself, spread over four different offices between which cooperation is not always optimal, and there is some duplication and disagreement. In particular, among the agents involved, there is some variation of *attitude* towards the norms mentioned, with different priorities apparently depending at least partly on personal preferences.

5. Concluding comments

I have argued that the literal translation hypothesis is an important one. It has interesting theoretical implications about the cognitive relation between form and meaning; it is testable in multiple ways; it has explanatory power, in that it connects with different kinds of research results and relates to several other claims, some of which run counter to it; and it opens up new research questions. In the wider context of revision in general, one wonders for example whether there might be a pattern: deliteralization in self-translation, and reliteralization when the revision is done by someone else. Or would the different tendencies correlate rather to text type? To what extent might the translator's knowledge of an institution's revision system affect initial textual choices and strategies? Would the knowledge of the coming revision encourage more risk-taking at the initial stage, for instance? Such new questions might lead to interesting new hypotheses.

Suppose the hypothesis turns out to be largely supported by empirical evidence (as now seems), but not always. Then, as I have been suggesting at different points, we could ask:

- Under what conditions (textual, situational, psychological, etc.) do we find in self-revision a tendency towards deliteralization, and under what conditions one towards reliteralization?
- Under what conditions do we find in other-revision each of these tendencies?
- If the conditions affecting deliteralization or reliteralization in these two forms of revision turn out to be similar, what might this imply about the most likely cause of the tendency?
- To what extent might (unconscious) cognitive factors be relevant, and to what extent (potentially conscious) attitudinal factors?
- How then could this information be best exploited to optimize translation quality control systems?

There are thus implications here for making further connections between the translation act (cognitive, psychological) and the wider translation event (situational, institutional, sociological).

References

Catford, J. C. 1965. *A Linguistic Theory of Translation*. Oxford: OUP.

Chesterman, A. 2007. "What is a unique item?" In *Doubts and Directions in Translation Studies,* Y. Gambier, M. Schlesinger and R. Stolze (eds), 3–13. Amsterdam/Philadelphia: John Benjamins.

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process*. Amsterdam/Philadelphia: John Benjamins.

Halverson, S. 2003. "The cognitive basis of translation universals." *Target* 15 (2): 197–241.

———. 2007. "Investigating gravitational pull in translation: the case of the English progressive construction." In *Texts, Process and Corpora: Research Inspired by Sonja Tirkkonen-Condit*, R. Jääskeläinen, T. Puurtinen and H. Stotesbury (eds), 175–195. Joensuu: Publications of the Savonlinna School of Translation Studies 5.

Ivir, V. 1981. "Formal correspondence vs. translation equivalence revisited." In *Theory of Translation and Intercultural Relations* [Poetics Today 2:4], I. Even-Zohar and G. Toury (eds), 51–59. Tel Aviv: Porter Institute for Poetics and Semiotics, Tel Aviv University.

Jakobsen, A. L. 1999. "Logging target text production with *Translog*." In *Probing the Process in Translation: Methods and Results*, G. Hansen (ed.), 9–20. Copenhagen: Samfundslitteratur.

Jansen, A. and Jakobsen. A. L. 2000. "Translating under time pressure." In *Translation in Context*, A. Chesterman, N. Gallardo San Salvador and Y. Gambier (eds), 105–116. Amsterdam/Philadelphia: John Benjamins.

Krzeszowski, T. P. 1990. *Contrasting Languages. The Scope of Contrastive Linguistics*. Berlin: Mouton de Gruyter.

Leuven-Zwart, K. M. van. 1989/1990. "Translation and original. Similarities and dissimilarities, I and II." *Target* 1 (2): 151–181 and 2 (1): 69–95.

Matthiessen, C. M. I. M. 2001. "The environment of translation." In *Exploring Translation and Multilingual Text Production*, E. Steiner and C. Yallop (eds), 41–124. Berlin/New York: Mouton de Gruyter.

Mauranen, A. and Kujamäki, P. (eds). 2004. *Translation Universals. Do They Exist?* Amsterdam/Philadelphia: John Benjamins.

Mossop, B. 2000. "The workplace procedure of professional translators." In *Translation in Context*, A. Chesterman, N. Gallardo San Salvador and Y. Gambier (eds), 39–48. Amsterdam/Philadelphia: John Benjamins.

———. 2007. *Revising and editing for translators* (Second edition). Manchester: St. Jerome Publishing.

Murphy, S. 2003. "Second language transfer during third language acquisition." *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics* 3, No. 2. http://journals.tc-library.org/index.php/tesol/issue/view/6. (Accessed January 7, 2007.)

Newmark, P. 1988. *A Textbook of Translation*. Hemel Hempstead: Prentice Hall.

Nida, E. A. 1964. *Toward a Science of Translating*. Leiden: Brill.

Nordman, L. 2009. *Lagöversättning som process och produkt. Revideringar och institutioner vid lagöversättning till svenska i Finland*. Helsinki: Department of Scandinavian Languages and Scandinavian Literature.

Pym, A. 2009. "Using process studies in translator training. Self-discovery through lousy experiments." In *Methodology, Technology and Innovation in Translation Process Research*, S. Göpferich, F. Alves and I. M. Mees (eds), 135–156. Copenhagen: Samfundslitteratur.

———. 2010. *Exploring Translation Theories*. London: Routledge.

Seleskovitch, D. and Lederer, M. 1984. *Interpréter pour traduire*. Paris: Didier.

Shuttleworth, M. and Cowie, M. 1997. *Dictionary of Translation Studies*. Manchester: St. Jerome Publishing.

Tirkkonen-Condit, S. 2004. "Unique items – over- or under-represented in translated language?" In *Translation Universals. Do They Exist?* A. Mauranen and P. Kujamäki (eds), 177–184. Amsterdam/Philadelphia: John Benjamins.

———. 2005. "The monitor model revisited: Evidence from process research." *Meta* 50 (2): 405–414.

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam/Philadelphia: John Benjamins.

Vinay, J.-P. and Darbelnet, J. 1958. *Stylistique comparée du français et de l'anglais*. Paris: Didier. Translated into English by J. Sager & M. J. Hamel, 1995; *Comparative Stylistics of French and English. A Methodology for Translation*. Amsterdam/Philadelphia: John Benjamins.

Wilson, A. 2009. *Translators on Translating. Inside the Invisible Art*. Vancouver: CCSP Press.

Tracking translators' keystrokes and eye movements with Translog*

Arnt Lykke Jakobsen
Copenhagen Business School, Denmark

Although keylogging opens up many possibilities for studying translation processes, the addition of eye tracking vastly increases our ability to know what kind of process a translator engages in at any given point in time. Following a description of how the Translog software was technologically reshaped in the context of the EU Eye-to-IT project to be able to integrate keylogging and eye tracking, and of how raw eye and key data were represented, an analysis of a small sample of eye and key data results will be presented in a tentative formulation of a recurrent six-step processing micro-cycle which introduces the "anchor" word as a novel concept. It is optimistically suggested that an elementary cycle of the kind suggested could both function as the basis of computational analysis of large volumes of translational eye and key data and potentially as the core of a computational model of human translation.

Keywords: translation processes, keylogging, eye tracking, Translog

1. Introduction

The Translog software was originally only developed to record translators' keystrokes in time. A complete, timed sequence of keystrokes provides a detailed overview of the entire typing process by which a translation comes into existence. Data records of this type make it possible to study cognitive phenomena such as chunking, based on the distribution of pauses, and such writing-related phenomena as edits and corrections. However, although it is sometimes possible to connect production data with comprehension phenomena, in most instances it is only possible to speculate on what comprehension processes a translator engaged in before a solution was typed, since the only recorded data is a pause.

---

* Very constructive comments from two anonymous reviewers and the editors of the volume are gratefully acknowledged.

The addition of eye tracking to keylogging adds very rich information on what goes on when no typing activity is recorded. Source-text reading can be tracked as well as target-text reading and monitoring, and also shifts of attention from source to target text and back. Such information radically improves our chances of reconstructing both the comprehension processes that precede production and the way in which comprehension and production processes combine.

Eye movement recording has made us aware of the magnitude of the work performed by the eyes – and the human brain – in reading. This has been described in several major studies of reading (Rayner and Pollatsek 1989; Rayner 1998; Radach, Kennedy and Rayner 2004), but the kind of reading that is part of written translation has not been studied. When we recorded the eye movements of participants instructed in experiments to read a foreign-language text for comprehension, and compared them to eye movements of participants who translated a screen text in the foreign language and typed the translation on the same screen, we found that written translation involves vastly more eye work and appears to be a considerably more demanding task, cognitively, than just reading and comprehending a text or even reading it and translating it orally (Jakobsen and Jensen 2008). Eye-movement observation helps us better understand the mechanical complexity of the task of written translation, but gaze data also help us penetrate closer to the core of the cognitive challenges in translation. The fundamental assumption here is that there is a correlation between behavioural "outside" data and cognitive "inside" processing. Though this so-called eye-mind hypothesis (Just and Carpenter 1980), which states that what the eyes are looking at is what the mind is attending to, has been demonstrated not to apply everywhere during a reading process (cf. e.g. Hyönä, Lorch and Rinck 2003), eye movement data still provide a fine window to the mind. By visualising the way in which reading proceeds and how it interacts with typing, we get the best available picture of how translation processes are coordinated in the mind.

2. Analysing keystroke data

The original impetus for the development of the Translog software was an ambition to supplement think-aloud process data from translation experiments with "harder", machine-recorded evidence of human translation processes. The methodological assumption behind this ambition was a conviction that if it were possible to record converging qualitative think-aloud data and quantitative behavioural data, our hypotheses about translation processes would gain important support.

The first Translog version, developed in 1995, had three main functions. It could display source text automatically as full text, in paragraphs, in sentences, or in segments defined by the researcher. The central logging function, in the component then called Writelog (Jakobsen and Schou 1999), recorded all keystrokes, including navigation and deletion keystrokes, but not yet mouse clicks, that were made during the

completion of a translation task in time, and then saved this data in a log file. Subsequently, information stored in the log file could be displayed in two modes. It could be replayed dynamically, and it could be represented as a linear sequence of keystrokes.

The replay could be done at different speeds, which allowed the researcher to inspect the temporal dynamics of text production in translation visually when replayed at normal speed, or somewhat more slowly or faster. Faster replay provided direct visual evidence of the chunked structure of translational text production. Slowed down replay had an observational effect similar to that of a magnifying glass. In our research group, we soon found that by deploying the replay function to support participants' recall during retrospection, we could obtain very rich data. When participants were given the chance to observe a replay of their keystrokes as they were being interviewed, they appeared to immediately and very exactly recall what had been on their mind at a certain point in time during a translation task.[1] Though it could not be known with any certainty that what was reported in retrospection was as reliable as concurrent verbalisation, the data elicited in this manner had all the appearance of truthful evidence. It was generally very detailed, and it was articulated spontaneously and with great conviction by participants.

One great advantage of retrospective data, in addition to the data being both rich and accurate, at least apparently, was that there could be no suspicion of such data impacting on the primary (keystroke) data since it was all elicited after the event. Ericsson and Simon's (1984) claim, that concurrent think-aloud had no effect on primary processing if verbalised content was available in STM in verbal form, was not supported by our findings in experiments in Copenhagen, which indicated that concurrent think-aloud had a degenerative effect on processing (Jakobsen 2003). For such reasons more and more experiments were run with cued retrospection rather than concurrent verbalisation (e.g. Hansen 2005 and 2006).

Towards the end of 1999, Lasse Schou finished programming the first Windows (98) version of Translog, which we named Translog2000. Early experimental versions of this new development had been available for testing and use in the CBS TRAP project (1996–2000). The three basic functions were still the same: a Translog project created in the Translog Supervisor module (formerly Translog) was displayed in Translog User (formerly Writelog), which logged the key and time data and saved binary data in a log file that could be inspected (replayed and represented linearly) in the Supervisor module, but was not otherwise immediately accessible. The program included a small statistical component, which calculated e.g. task time and the total number of keystrokes.

Two of the main applications of this version of Translog were to the study of translation phases and to how translation students and professional or expert translators

1. There is a fine description of the method in Englund Dimitrova (2005: 66) and an extended discussion of retrospection as a research method in Englund Dimitrova and Tiselius (2009).

chunked their typing of a translation. In the context of the Expertise project[2] several experiments demonstrated that professional or expert translators spent more time and effort on checking and revising their translation than did translation students, even though their translation was already of superior quality. It was also demonstrated clearly that expert translators chunked their production into longer segments than did students. This strongly suggested that expert translators handled the translation task cognitively differently in that they processed longer strings of text and longer translation units than did student translators. Comparison of segment length in the initial portions of the drafting phase with segment length further on in the drafting phase gave quantitative empirical support for the facilitation effects hypothesis.[3]

The fact that Translog provided a complete record of all revision acts made it suitable for studying editorial practices, e.g. the distribution of "online" revision done during the drafting phase and end revision or for more qualitative studies of revision. The software was also applied to the study of various distracting phenomena in the production of translations (Hansen 2005).

Although keystroke logging provides data for important insights into the size and composition of the chunks from which translators construct their target texts, it is very difficult to know from a keystroke record what cognitive activity was going on in the pauses between chunks, why pauses had the duration they did, or why they occurred. In most of the recordings we made, chunks were relatively short with fewer than four or five words and with pauses of about two seconds or more between them. This seemed to suggest a process which alternated between source-text reading and comprehension, mental formulation of the translation and physical typing of the translation. However, in some instances the chunks we recorded were so long that they far exceeded anything that could be held in short-term memory at any one time (so-called instances of peak performance, cf. Jakobsen 2005) and were not explicable in terms of a serial succession of comprehension, formulation of the translation and typing of it. Occurrences of that kind seemed to indicate that at least some translators were able to continuously feed their short-term memory and process the information concurrently with their typing of the translation. Such evidence of overlapping comprehension, formulation and typing activity was strongly suggestive of parallel processing.

But this was only one problem that was in obvious need of supplementary evidence. Fundamentally, we could only speculate from the linguistic environment in

---

2. Details at http://www.hf.uio.no/ilos/forskning/forskningsprosjekter/expertise/.

3. The term "facilitation effects" has been used in the study of L2 acquisition and bilingualism since the 1960s to indicate features in the L1 that facilitate the acquisition of L2. In translation research the term refers to the observation that translation of a text gets easier and faster as the translator gets into the text. See Englund Dimitrova (2005: 98): "The possible facilitation effect during the writing phase of the translation task will also be investigated, more particularly whether there are differences in segmentation and/or number of characters per segment at the beginning and towards the end of the writing phase of the task."

which pauses occurred why they occurred and what processing they might be correlated with. The data, coming at the tail end of the translation process, was open to multiple interpretations. We could not even know if pauses were always caused by task-related phenomena or by factors totally irrelevant to the execution of the task. An obvious way to look for supplementary evidence was by trying to establish what translators were looking at during a translation task.

3. Presumed research benefits from adding gaze data to keystroke data

One of the fundamental ambitions in the EU Eye-to-IT project (2006–2009)[4] was to develop an application that would integrate reading data ("gaze" or "eye" data) from an eye tracker with the kind of writing data that was already available in programs like Scriptlog, Inputlog and Translog.[5] The challenge was to map information from the eye tracker to the words a translator actually reads.

A fundamental assumption was that eye data would provide evidence pertaining identifiably to source-text reading so that source-text comprehension processes could be studied separately from text-production processes and could be compared with other reading processes that were not part of a translation process. Eye data were expected to be available all through a translation task without interruption or with few interruptions and to be able to "fill" most of the pauses in keystroke activity with interesting data.

Another assumption was that integration of eye and key data would provide new opportunities for researching the dynamics of how source-text comprehension activity, reflected in translators' eye movements, was coordinated with target-text production activity, reflected primarily in keystroke activity but possibly also in gaze activity.

A further assumption was that just as the distribution of pause duration in the writing process seemed relatable to translation difficulty or ease, so fixation duration and eye movement patterns, such as regressions, would be correlatable with comprehension difficulty or ease as well as with text production and/or translation difficulty or ease.

Finally, it was assumed that if it were possible to map the eye tracker's gaze data reliably to the words being read on the screen in real time, i.e. if the eye tracker's gaze data (giving the pixel position of the gaze) could be instantly "translated" into information about what words were being looked at, we would be able to exploit such real-time mapping to support translation with prompts activated by the translator's gaze.

4. See http://cogs.nbu.bg/eye-to-it/.

5. See http://scriptlog.net/, http://webh01.ua.ac.be/mleijten/inputlog/ and http://translog.dk.

3.1 Translog development in the Eye-to-IT project

It soon became clear that Translog2000 would be difficult to update in a way that would meet the promises made in the Eye-to-IT research description. Therefore it was decided that Lasse Schou would completely rewrite the program in C#, using Microsoft's .NET 2.0 environment, and introducing Unicode 8 (to accommodate non-Latin scripts) and xml as an open output format. The result was Translog2006, currently distributed as the "academic" version. (The distributed version does not include the gaze recording component that was developed experimentally for the Eye-to-IT project.[6])

The big technical challenges in the Eye-to-IT project were to find a way of integrating keystroke and gaze recording so that it would be possible to have a perfectly timed sequence of eye and key data, and to find a way of mapping information from the eye tracker in real time to what words were currently being read by a user. This dual technical and programming challenge was taken on jointly by Lasse Schou and Oleg Špakov, and in the course of the EU Eye-to-IT project (2006–2009) a gaze-to-word mapping application (GWM), developed at the University of Tampere,[7] was successfully integrated into an experimental version of Translog. Gaze samples from the Tobii 1750 eye tracker were interpreted as saccades or fixations, and the GWM module automatically mapped fixations to a source-text or target-text word on the basis of the current positions of words on the screen, with some adjustment performed by an algorithm representing a crude model of reading. Raw gaze data were saved in separate files, and only mapped, interpreted data (the succession of fixations with specified durations, mapped to words) were recorded in Translog. The output from the integrated Translog/GWM application was a perfectly timed record of the sequence of a user's word fixations across the screen and the user's keystrokes (cf. Figures 2 and 3). In Translog, the data could be represented in linear fashion with information at millisecond level about the intervals between gaze or keystroke events, and data could also be replayed with keystrokes appearing dynamically as in earlier Translog versions, but now also dynamically highlighting the user's recorded fixations on source or target text words for as long as a fixation on a word lasted (see Figure 1).

In earlier versions of Translog raw data were recorded in binary code, but Translog2006 stores all raw data in openly accessible xml code. Figures 2 and 3 show an extract of the data underlying the screen representation in Figure 1. Figure 2 gives information about what kind of "event" (fixation (7), keystroke (1), synchronisation signal (6)) was logged; the Value parameter indicates what event subtype occurred

---

6. A version which includes keystroke, gaze and audio recording as well as real-time gaze-to-word mapping is currently being tested. Hopefully, it will be ready for release in the course of 2011.

7. The solution was inspired by the iDict project developed at the University of Tampere (Hyrskykari, Majaranta, Aaltonen and Räihä 2000; Hyrskykari 2006).

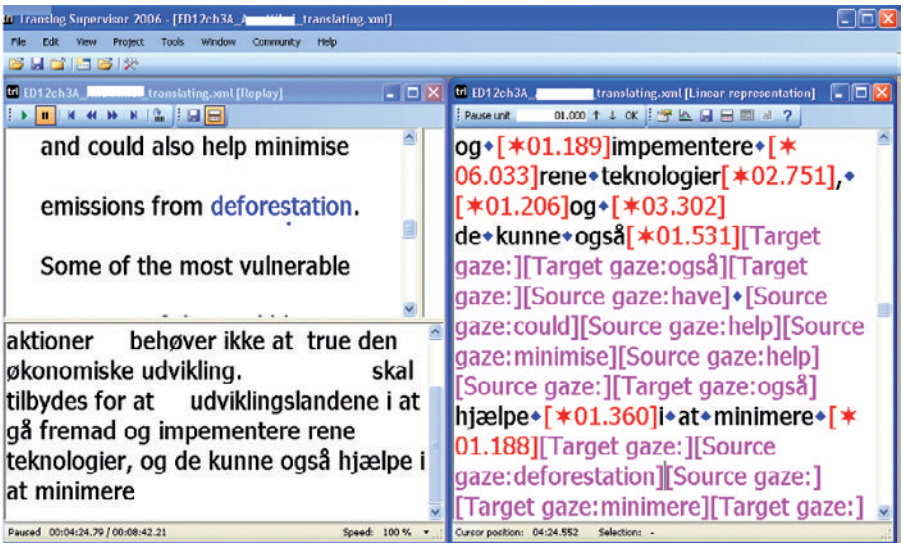


Figure 1. Translog2006 display of source text in the upper left window and an extract from the linear representation of gaze and keystroke data in the right-hand window, showing the current cursor position after [source gaze: deforestation]. When activated, the replay displays the succession of keystrokes in the lower left window, while highlighting the current source-text word ['deforestation'] looked at, at the replay speed selected. Source text and (more often) target text fixations sometimes failed to be mapped [Target gaze:]. The occurrence of pauses lasting more than a second where no data was recorded, e.g. [*02.751], indicates that the translator was looking away from the screen, most probably to attend visually to the keyboard while typing.

```
<LogEvent Action="7" Value="2" Cursor="295" Block="4" Time="00:04:15:310" />
<LogEvent Action="7" Value="2" Cursor="-1" Block="0" Time="00:04:15:437" />
<LogEvent Action="7" Value="1" Cursor="438" Block="4" Time="00:04:15:542" />
<LogEvent Action="1" Value="32" Cursor="300" Block="0" Time="00:04:15:544" />
<LogEvent Action="7" Value="1" Cursor="332" Block="5" Time="00:04:15:762" />
<LogEvent Action="7" Value="1" Cursor="343" Block="4" Time="00:04:15:951" />
<LogEvent Action="7" Value="1" Cursor="348" Block="8" Time="00:04:16:136" />
<LogEvent Action="7" Value="1" Cursor="343" Block="4" Time="00:04:16:548" />
<LogEvent Action="7" Value="1" Cursor="-1" Block="0" Time="00:04:16:695" />
<LogEvent Action="7" Value="2" Cursor="295" Block="4" Time="00:04:17:100" />
<LogEvent Action="1" Value="104" Cursor="301" Block="0" Time="00:04:17:420" />
<LogEvent Action="1" Value="106" Cursor="302" Block="0" Time="00:04:17:815" />
<LogEvent Action="1" Value="230" Cursor="303" Block="0" Time="00:04:18:247" />
<LogEvent Action="1" Value="108" Cursor="304" Block="0" Time="00:04:18:447" />
<LogEvent Action="1" Value="112" Cursor="305" Block="0" Time="00:04:18:627" />
<LogEvent Action="1" Value="101" Cursor="306" Block="0" Time="00:04:18:715" />
<LogEvent Action="1" Value="32" Cursor="307" Block="0" Time="00:04:18:832" />
<LogEvent Action="6" Value="1" Cursor="0" Block="0" Time="00:04:20:192" />
```

Figure 2. Extract of the xml record of some of the raw data underlying the representation shown in Figure 1.

```
<string>også</string>
<string xsi:nil="true" />
<string>have</string>
<string>could</string>
<string>help</string>
<string>minimise</string>
<string>help</string>
<string xsi:nil="true" />
<string>også</string>
<string xsi:nil="true" />
<string>deforestation</string>
<string xsi:nil="true" />
<string>minimere</string>
<string xsi:nil="true" />
<string>minimise</string>
<string>emissions</string>
```

Figure 3. Automatically generated list of successively fixated words in the source and target texts. Note that in five instances a fixation was recorded, but word mapping failed (xsi:nil="true").

(source-text fixation (1), target-text fixation (2); keystroke on the space bar (32), letter h (104), etc.). The Cursor and Block parameters show where in the sequence of events a given event occurred, and how many positions the event moved the cursor, and Time gives the time (hrs:mins:secs:mss) at which an event occurred, starting at the moment when the task was begun.

In the first line of Figure 2, <LogEvent Action="7" Value="2" …Block="4"…> indicates that the LogEvent was a gaze event ("7") in the target text window ("2"), where the word looked at ["også"] had 4 characters (Block="4"). <LogEvent Action="7" Value="1"…> in the third line represents a fixation on a word ["have"] in the source text, with four letters. <LogEvent Action="1" Value="32"…> in line 4 indicates that the key with ASCII/Unicode number 32, [<space>], was struck, and that the space was inserted at position 300 at the time indicated. Then follow five fixations on words in the source text ["could", "help", "minimise", "help", Ø] and one in the target text ["også"], after which we get a sequence of keystrokes making up the target text "hjælpe". (Log event 6 indicates a synchronisation signal sent by Translog.)

The words to which fixations are automatically mapped by the system are listed elsewhere in the xml log file. If mapping is complete and successful, every fixation will be mapped to the word that was being looked at, at the time the event was recorded. Figure 3 shows an extract from such a list.

4. The analytical challenge: Preliminary process observations

With recorded data organised in this structured manner, we face a new analytical challenge. All the details in a translator's eye-key activity can be inspected, but to begin to make coherent sense, details have to be interpretable as constituents in a model. The translator recorded in Figures 1–3 spent considerable time attending visually to the keyboard during typing. For this reason, there are long periods of more than 10 seconds with no recorded eye data in this particular recording. Another consequence of this is that we get fairly extended periods of typing without intervening visual attention to the source text. Finally, because of the extended visual attention to the keyboard, which is cognitively much like a third text, the gaze data show that it often took this translator some rereading to get back to the next relevant source-text segment.

Eye-movement recording visualises how much work is performed by the eyes – and the human brain – in reading. The challenges, from the point of view of translation process research, are to investigate how reading, when part of a process of translation, differs from other kinds of reading, and how reading interacts with other cognitive processes involved in translation. Our findings in 2007, when we recorded the eye movements of participants instructed in experiments to read a foreign-language text for comprehension, and then compared them to eye movements of participants who translated a screen text in the foreign language and typed the translation on the same screen, showed that eye movements in reading are highly dependent both on the purpose of reading (cf. early observations by Yarbus 1967) and on the nature of the reading task. Written translation involves vastly more eye work and appears to be a considerably more demanding task, cognitively, than just reading and comprehending a text or even reading it and translating it orally (Jakobsen and Jensen 2008). This may be in part because the modality of speech has been better automatised than the modality of writing, but it also appears to be caused to some extent by the necessary ocular motor effort involved in helping the brain keep track of what source-text words can be or have been aligned with what target-text words. Alignment work forces the brain to constantly instruct the eyes to shift between the source and target texts. Though eye tracking provides ample evidence that many translators have very accurate spatial text memory and can shift between local text strings in different texts with great accuracy, such transitions may also cause confusion and make reorientation and rereading necessary. Such random "noise" in the data makes it more difficult to detect regularities.

A preliminary report of the shifts of attention between source and target texts can be generated from lists such as the ones shown in Figures 2 and 3, and the points at which text was typed can also be reconstructed. Such a report reveals that the eyes are constantly engaged in three basic types of activity: (a) source-text (ST) reading, (b) target-text (TT) reading and (c) monitoring of the keystroke activity and outcome. We know that translation must also be taking place, most probably in parallel with other processes, but behavioural evidence of this activity is only indirect. At the time recorded in Figures 1–3, the translator had just typed "og de kunne også"

(a translation of "and could also"). The ST text fragment processed in the extract shown in Figure 3 was "help minimise". Looking more closely at the sequence of word fixations, we make the following observations and interpretations: (1) The most recently typed word ("også") was fixated immediately after having been typed (an example of TT production being monitored). (2) The fixation of "have", which appeared below the intended target word "help" one line down,[8] illustrates the phenomenon of "undershooting", known from studies of eye movements at line shifts in ordinary reading. The distance from "også" in the target text up to "help" in the source text required a saccade almost the length of that required at line shifts, and therefore the target word "help" was, apparently, not directly fixated. It is possible, however, that the translator's brain perceived the target word parafoveally because after the short stopover on "have", the saccade veered off to the left to "could", a word already translated. (3) As "could" had already been translated, a probable reason for rereading that more distant word, rather than going to the more proximate new ST target word ("help"), could have been the need to establish a fresh mental anchor word for the next chunk of source text to be processed. (4) Having established the anchor, the mental ground had been prepared for the next two target words, "help" and "minimise", to be successively fixated. (5) Though one might expect the translator to have wanted to fixate the next word also ("emissions", the object of the transitive verb "minimise"), this did not happen until later, perhaps an indication that this translator did not engage in very deep parsing of the source text. Instead, this translator's gaze regressed to briefly refixate "help" (and perhaps "minimise") while probably formulating a translation. (6) Before typing the translation, however, the last word of the current target text was fixated, establishing a target anchor for the translation of "help minimise" to attach to. (7) Only then did the translator type the (unidiomatic) translation "hjælpe i at minimere", without looking at the screen. (8) Finally, the new target text screen input was monitored (in an unmapped fixation). Having performed this sequence of actions, the translator was ready to embark on translating the next chunk, "emissions from deforestation", in much the same manner.

5. A recurrent processing pattern

Eye movement observation helps us better understand the mechanical complexity of the task of written translation, but gaze data also help us penetrate closer to the core of the cognitive processes in translation. The eye-mind hypothesis (Just and Carpenter 1980) is the fundamental assumption that what the eyes are looking at is what the mind is attending to. The correlation has been demonstrated not to apply universally during a reading process (cf. e.g. Hyönä, Lorch Jr. and Rinck 2003), but although the correlation is not always perfectly straightforward, the basic assumption still stands

8. Figure 4 shows the source text as laid out on the screen in the experiment.

that what the eye is looking at is (in general) something the mind is attending to. So eye movement data provide a fine window to the mind, and by visualising the way in which reading proceeds and how reading interacts with typing, we get the best available picture of how translation processes are coordinated in the mind.

With a structured record of the exact temporal succession of translators' eye and hand activity, we can begin to ask old questions in a new way and begin to formulate tentative answers to them, as was attempted in the previous section.

One such question concerns the way in which translators segment their production of the target text. Instead of basing our account of segmentation on temporal aspects of text production (alternating sequences of pauses and production periods), we can now enrich our analysis with eye data. As we have just seen, eye data provides strong evidence of how much source text was read prior to the production of a piece of target text. This gives us an exact idea of what source text segment was read immediately prior to the typing of the matching piece of target text. Instead of mechanically defining a segment or a translation unit by the occurrence of a pause in the typing activity of a certain duration, we can now more qualitatively, and with greater detail, identify what source text segment was being read and what corresponding target text was being typed, regardless of the duration of any segment boundary pauses.

Another major question concerns the coordination of comprehension, translation and formulation processes. The addition of gaze data gives us a finer-grained picture of the behavioural and cognitive complexity of the eye-hand (and brain) interaction that goes on in translation than we had before. Though we still only get indirect evidence of translation processing going on, we get stronger evidence of the steps that the human translator's brain goes through to compute the trajectory of the eyes and the finger movements across the keyboard in the execution of an act of translation.

The observations and interpretations made in the previous section can be summarised schematically. In Table 1, the actions have been tabulated to give a survey of the sequence of actions and outcomes.

Table 1. Visual attention shifts between ST and TT based on automatic mapping of gaze data to words looked at. The TYP column indicates where typing was done. After the second fixation of "også", the words "hjælpe i at minimere" were typed. Question marks indicate unmapped fixations.

TT	ST	TT	TYP	TT	ST	TT	ST
	have				deforestation		Minimise
	could				(?)		emissions
	help		hjælpe				
	minimise		i				
	help		at minimere				
	(?)						
også		også		(?)		minimere	
(?)						(?)	

The sequence of actions performed by our human translator can now be hypothetically abstracted into a small algorithm. Starting with an act of comprehension, the cycle of cognitive and motor processes that we can establish on the basis of the above example data involves the following six steps:

1. Moving the gaze to read the next chunk of new source text (and constructing a translation of it),
2. Shifting the gaze to the target text to locate the input area and read the current target-text anchor word(s),
3. Typing the translation of the source-text chunk,
4. Monitoring the typing process and the screen outcome,
5. Shifting the gaze to the source text to locate the relevant reading area,
6. Reading the current source-text anchor word(s).

These steps should be seen as constituting a complete micro-cycle in which some of the steps (or portions of them) can be skipped or, conversely, can be repeated several times. For instance, in step 2, the gaze will be shifted to the target text, but not necessarily to an anchor word. We also frequently find that the gaze may be shifted back and forth several times between the source text word(s) being processed (step 2) and the translation being typed (steps 3 and 4). A finer-grained model would therefore add subcategories to each step or some of them and would point out that e.g. steps 3 and 4 can be executed quite differently, depending on the translator's typing skills. It should also be noted that steps 3 and 4 are to some extent concurrent. Monitoring takes place concurrently with text being typed, but (screen) monitoring frequently continues after typing has stopped. However, even without such detail, the model allows us to predict with a high degree of probability that step 1 will be followed by step 2, step 2 by step 3, etc.

As can be seen, the ocular motor effort involved in helping the brain keep track of what source-text words have been read and translated, of what words can be or have already been aligned with what target-text words, and of what target-text words have been typed, and how, is very considerable. Alignment work ("vertical" alignment) across the two languages involved forces the brain to constantly instruct the eyes to shift between the source and target texts. Monolingual ("horizontal") syntagmatic attachment also necessitates rereading when new text is added to old text, visually or typographically.

6. Sample data analysis

The availability of large volumes of structured data makes it obvious to look for support for observations we make on the basis of small data samples from applications of statistical methods to big data corpora. To do this, we need to be sure that the data collected satisfies basic empirical criteria, e.g. regarding completeness. As is evident from the sample shown (see esp. Figure 3), the current system does not always

succeed in mapping fixations. A more fundamental challenge concerns the discovery of fixations. This is at once a technical and a theoretical challenge. The eye tracker that was used in our experiments was a 50 Hz machine, i.e. it recorded gaze samples at 20 ms intervals. With such equipment, the traditional method of identifying fixations is to say that a fixation occurred if five consecutive samples were recorded within a specified area, e.g. no more than 40 pixels apart. This means that a fixation is always recorded as having a duration of at least 100 ms. Though more sophisticated, velocity-based definitions are now gaining currency, this does not change the fact that what constitutes a fixation is highly dependent on the equipment used. It would be absurd, for instance, to use the traditional definition on data from a 1000 Hz eye tracker and claim that a fixation could be as brief as 5 ms, but it might make perfect sense to claim that fixations can occasionally have shorter durations than 100 ms. There is also a possible upper limit that has to be considered. We need to consider how long a fixation can be before we need to split it into several fixations. There are several other technical challenges to getting good reading data from translation experiments, especially because of the frequent shifts of visual attention and the discontinuous nature of the reading. In addition, gaze data recording is sensitive to a large number of variables, including head movement, calibration accuracy, blinks and the shape of the translator's eye lashes. With poor-quality data, the danger is that fixations, however defined, are not even discovered. And obviously, if fixations are not discovered, there is nothing to be mapped.

To illustrate the variability in the different data records we are challenged by, Figure 4 shows the succession of fixations (still in the same extract) as shown in another

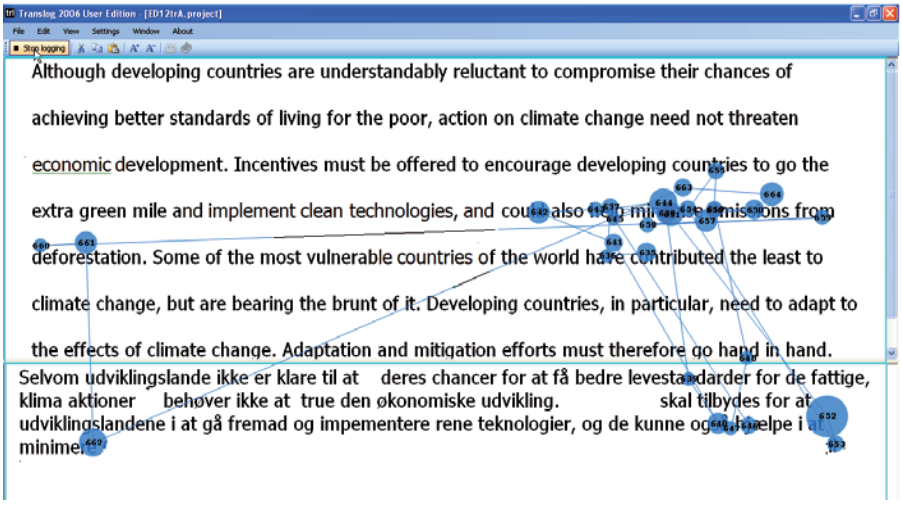


Figure 4. Extract from a ClearView gaze plot representation of fixations (635–664). Fixations 640–664 were recorded in the exact same time span and from the same event as that recorded in Figures 2–3. (Fixations 635–639 were earlier.)

software program. Figure 4 shows a gaze plot generated by the Tobii company's ClearView software, which records the succession of fixations and saccades and plots them to the computed, approximate position on the screen, with no adjustment to what is displayed on the screen. When the gaze plot is laid over the text that was displayed when the recording was made, it can be seen that according to ClearView, there were probably several more word fixations than the eleven fixations that were successfully mapped by the GWM module reporting to Translog (cf. Figure 3).

The representation of fixations in Figure 4 illustrates some of the theoretical and technical problems involved in connecting recorded fixations reliably to what we think was probably seen and read. These problems go to the heart of the eye-mind relationship: How do the fixations computed by our equipment correlate with the mental experience of reading a sequence of words? Our equipment records what the physical eye appears to be targeting, but exactly how this recording corresponds to what the mind sees is a major issue. Many of the fixations are not recorded as having been on a specific word, but appear slightly below or above a word, or a fixation is recorded as being between words or at one extreme of a word, or else two (or more) successive fixations are recorded within the same word. Manual mapping of such data is a very instructive reminder that what we call a fixation is neither a very well-defined concept nor an easily identified object, but a technology and theory-dependent abstraction with a certain grounding in behaviour.

One way of making sense of the gaze data represented in Figure 4 is by interpreting it in the light of a model of reading or, better, in the light of a process model of translation which includes a model of probable reading behaviour. The function of such a model is to close the gap, as far as possible, between the recorded fixations and the words that were actually read by the participant. Table 2 presents the result of such a process. Fixations 640 through 647 could be mapped identically without any problems. (Translog/GWM interpreted fixation 640 as two separate fixations, but was unable to map the second. Fixation 646 looks like a fixation on "hjælpe" in ClearView, but since the timestamp information shows that the word had not yet been typed at the time of the fixation, it had to be mapped to "også".) Fixations 660–664 were also mapped almost identically (Translog/GWM collapsed 660 and 661 into a single fixation on "deforestation"). Fixation 663 was recorded in ClearView as being somewhere between "minimise" (in line 4) and "countries" (in line 3). A decision based on proximity would align this fixation with "countries", but this mapping conflicts with the 6-step model outlined above, which predicts that a fixation following a fixation that serves to monitor target output (step 4) will be followed by a shift of the gaze to the ST (step 5), followed by a fixation on an anchor word (step 6). The expected trajectory would therefore be from "minimere" (target output monitored) to ST anchor "minimise", with perhaps a quick stopover near the next target word ("emissions") to adjust the trajectory. Therefore, the decision was to map this fixation to "minimise" in accordance with the model, and to similarly map fixation 664 to "emissions" rather than to "to go" in the line above.

Table 2. Automatically mapped fixations by Translog/GWM compared to manually mapped fixations from ClearView. Between fixations 647 and 648 there was an interval of 2.013 seconds in which the word "hjælpe" was typed. Between fixations 651 and 654 there was an interval of 3.054 s during which the words "i at minimere" were typed.

Fixated words	Time extract: 13 seconds, 509–521 seconds into task				
Translog	ClearView's record of fixations, manually mapped to words				
	fix	**ST word**	**TT word**	**Dur**	**Comment**
	635	Contributed		239	
	636	have		120	line below "help"
	637	help		179	
	638	minimise		239	
	639		levestandarder	100	on way down to TT "end"
også (+unmapped)	640		også	259	
have	641	have		199	line below "help"
could	642	could		299	
help	643	help		140	
minimise	644	minimise		399	
help	645	help		120	
(unmapped)	646		også	159	
også	647		også	120	
	648	hand		120	on way up to "emissions"
	649	emissions		179	first half of word fixated
	650	emissions		219	second half of word fixated
	651	minimise		319	
	652		at	698	
	653		at	179	
	654	minimise		179	
	655	countries		160	line above "emissions"
	656	emissions		140	first half of word fixated
	657	emissions		279	first half of word fixated
	658	minimise		139	
	659	from		140	
deforestation	660	defo(restation)		120	
	661	(defo)restation		279	
minimere	662	minimere		419	
minimise	663	minimise		199	
emissions	664	emissions		321	

Thus, with respect to fixations 640–647 and 660–664, there is almost perfect agreement between the automatic mapping result from Translog/GWM and from the

manual mapping of the ClearView gaze plot representation of recorded data. However, the data do not always support identical interpretations. ClearView fixations 648–659, for instance, were not mapped by Translog/GWM. They were not even identified as fixations. This may have been caused by the fact that between fixation 647 and 648 there was an interval of 2.013 s in which the translator typed the word "hjælpe" while probably looking at the keyboard. When the translator's gaze returned to the screen, GWM did not recognise the translator's gaze as quickly as ClearView did. Whatever the technical reason for this (perhaps some slight inconsistency between the ClearView and the GWM calibration data for this translator), missing data is a serious problem, but in this particular case we can fill the gap by manually mapping the fixations recorded in ClearView, as was done in Table 2. As we shall see, the data illustrates that the six steps in our model are not always followed strictly sequentially. Steps can be skipped or repeated several times, and occasionally processing may even backtrack to earlier steps if regular forward step-by-step processing fails to yield a relevant result. Fundamentally, however, we find the same basic processing cycle when we reconstruct the data missing in GWM. This is very good news for our chances of analysing the data automatically and for modelling the human translation process computationally. Let us take a closer look at the manually mapped sequence of fixations (643–660).

If we simplify the sequence slightly (collapsing multiple sequential fixations on the same word and disregarding the irregular fixation on "countries"), we get the following sequence of fixations: "hand", "emissions", "minimise", "at", "minimise", "emissions", "minimise", "from", "deforestation". In terms of our model, the sequence would be interpreted as representing the following steps: After fixating a succession of new source-text words, i.e. "help", "minimise" and "help" (step 1, fixations 643–645), the gaze was shifted to the target text, where fixations 646 and 647 identified the target text anchor word, "også" in two (repeated) fixations (step 2). This was followed by the typing of "hjælpe" (step 3), a translation of the first of the two new words read in 643–645. The absence of gaze data during the subsequent two seconds indicates that visual (monitoring) attention was on the keyboard only (step 4). After that, the gaze was shifted towards the next new piece of source text (step 5). Since the next new segment that was read ("minimise emissions") overlapped with the previous segment ("help minimise"), "minimise" functioned as the mental anchor (step 6) for the next new segment ("minimise emissions", step 1). This step was followed by a shift of the gaze to the target text area (step 2) and beyond that to the keyboard, as the translation of the rest of "minimise emissions" was typed ("i at minimere", step 3). The fact that no target-text anchor was fixated and the fact that two successive source-text segments overlapped suggest that "help minimise" was treated as one translation segment even if "emissions" was not fixated until after the translation of "help" had been typed. During the typing of "i at minimere" (step 3), which lasted 3.054 s, two fixations on "at" were recorded, indicating that visual attention was shifted from the keyboard to the screen (during step 4) for about one third (0.9 s.) of the time it took to type this portion of the translation. After the 3-second typing period, which might well indicate

cognitive uncertainty about the unidiomatic solution produced at this point, a new cycle was initiated, beginning with the shift of the gaze to the source text (step 5) and followed by fixation of the source-text anchor word "minimise" (step 6, fixation 654). This was followed by a reading of new source text ("emissions from deforestation"), but apparently this was too big a chunk for this translator to handle in a single processing cycle, so the anchor word "minimise" and its object "emissions" were refixated, after which the translator was ready to shift the gaze to the target text (step 2) and type the translation of "emissions" (step 3). As can be seen, the basic processing model is recognisable even in a sequence characterised by repeated fixations and backtracking. This gives the model potential for being used to handle fixation mapping problems and to diagnose e.g. parsing or meaning construction problems. It also opens up a new opportunity to computationally record the succession of new source-text segments read by a translator and to automatically align the source-text segments read and how they were translated.

7. Conclusion and perspective: Toward a computational model of human translation?

The addition of eye tracking to keylogging has made it possible to study what source or target text units are being worked on by a translator at any given point in time in much greater detail than has hitherto been possible. The data that we get from such integrated logging provides tangible evidence of the complex processing activity required by a translator. All along, the aim of the Translog development reported in the present article has been to enable us to know more certainly, from behavioural evidence, what the nature of this processing activity is, what steps are involved, what segments are read and aligned, and how this whole process is monitored. Our detailed analysis of a small sample of eye and key data showed that with the current state of the art there are still some technical issues that we should strive to improve. However, the analysis also indicated that there appears to be a micro-cycle of processing steps at the heart of translation. The integrated succession of eye and key data suggests that the basic dynamics of translational processing can be described in terms of six basic steps.

With the present state of technological development, it still seems relevant, perhaps even necessary, to examine small volumes of eye movement and keystroke data manually and selectively, both in order to measure the success rate of the system's automatic mapping and to develop hypotheses from small samples of carefully controlled, clean data that can be tested against bigger samples of somewhat noisier data. But this should not prevent us from pursuing a larger goal. The potential for large-scale computational analysis of translator activity data is there, or will very soon be there, and the prospect of creating a computational model of how expert human translators typically execute their skill seems within reach.

The eventual success of a computational model of human translation depends crucially both on the quality of the recorded raw data and on the soundness of our system's automatic interpretation of the data, primarily in the way fixations are identified and in the mapping algorithm. These interpretations will be informed initially by the knowledge of human translation that has been accumulated in the research community, but this knowledge might well be enriched by the discovery of new patterns and regularities in large volumes of data about the behaviour of human translators' hands and eyes.

References

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process.* Amsterdam/Philadelphia: John Benjamins.

Englund Dimitrova, B. and Tiselius, E. 2009. "Exploring retrospection as a research method for studying the translation process and the interpreting process." In *Methodology, Technology and Innovation in Translation Process Research*, I. Mees, F. Alves and S. Göpferich (eds), 109–134. Copenhagen: Samfundslitteratur.

Ericsson, K. A. and Simon, H. A. 1984. *Protocol Analysis. Verbal Reports as Data.* Cambridge, Mass.: MIT.

Hansen, G. 2005. "*Störquellen in Übersetzungsprozessen. Eine empirische Untersuchung von Zusammenhängen zwischen Profilen, Prozessen und Produkten.*" (Habilitation thesis. Copenhagen: Copenhagen Business School).

———. 2006. *Erfolgreich Übersetzen. Entdecken und Beheben von Störquellen.* Tübingen: Narr.

Hyönä, J., Lorch Jr., R. F. and Rinck, M. 2003. "Eye movement measures to study global text processing." In *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research,* J. Hyönä, R. Radach and H. Deubel (eds), 313–334. Amsterdam: Elsevier.

Hyrskykari, A. 2006. "*Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid.*" (Doctoral dissertation. [Dissertations in Interactive Technology, Number 4] University of Tampere, Finland. http://acta.uta.fi/pdf/951-44-6643-8.pdf.)

Hyrskykari, A., Majaranta, P., Aaltonen, A. and Räihä, K.-J. 2000. "Design issues of iDICT: A gaze-assisted translation aid." In *Eye Tracking Research & Application. Proceedings of the 2000 Symposium on Eye Tracking Research & Applications,* 9–14. New York: ACM.

Jakobsen, A. L. 2003. "Effects of think aloud on translation speed, revision and segmentation." In *Triangulating Translation. Perspectives in Process Oriented Research,* F. Alves (ed.), 69–95. Amsterdam/Philadelphia: John Benjamins.

———. 2005. "Investigating expert translators' processing knowledge." In *Knowledge Systems and Translation,* H. V. Dam, J. Engberg and H. Gerzymisch-Arbogast (eds), 173–189. Berlin/New York: De Gruyter.

Jakobsen, A. L. and Jensen, K. T. H. 2008. "Eye movement behaviour across four different types of reading task." In *Looking at Eyes. Eye-Tracking Studies of Reading and Translation Processing,* S. Göpferich, A. L. Jakobsen and I. M. Mees (eds), 103–124. Copenhagen: Samfundslitteratur.

Jakobsen, A. L. and Schou, L. 1999. "Translog Documentation." In *Probing the Process in Translation: Methods and Results,* G. Hansen (ed.), 151–186. Copenhagen: Samfundslitteratur.

Just, M. A. and Carpenter, P. A. 1980. "A theory of reading: from eye fixations to comprehension." *Psychological Review* 87: 329–354.

Radach, R., Kennedy, A. and Rayner, K. 2004. *Eye Movements and Information Processing During Reading.* Hove: Psychology Press.

Rayner, K. 1998. "Eye movements in reading and information processing: 20 years of research." *Psychological Bulletin* 124 (3): 372–422.

Rayner, K. and Pollatsek, A. 1989. *The Psychology of Reading*. Englewood Cliffs: Prentice Hall.

Yarbus, A. L. 1967. *Eye Movements and Vision*. New York: Plenum Press. (Translated from Russian by Basil Haigh. Original Russian edition published in Moscow in 1965.)

Seeing translation from inside the translator's mind

Brian Mossop
York University School of Translation, Toronto

A method is suggested whereby translators can state their beliefs about the meaning of the source text and their intentions about the wording of the translation. The method involves inserting items from a vocabulary into the translation as it is produced. The vocabulary consists of descriptors for various ways of producing language. Several examples are discussed: VENTRILOQUISING, QUOTING, FUDGING, GUESSING, EXPLICITATING, FICTIONISING, FABRICATING and GLOSSING. These words express states of mind of the translator which are not accessible to external observers. Some uses of the method are suggested.

Keywords: self-analysis, intentions, additions, connectors, uncertainty

1. Introduction

If we want to understand what is happening as a translator produces language, we need to know more than what external observation can tell us. As the early 20th-century Swiss linguist Charles Bally pointed out, there is a difference between people reflecting on their own acts and people's acts being the object of reflections by others:

> What others see of us are the external manifestations of our being, or rather, they see only the outcomes and symbols of our internal life in the form of actions, voluntary or involuntary movements, words, etc. These symbols may be rightly or wrongly interpreted, but even when the interpretation is correct and does indeed match our internal life, it is only a transposition of it; it is as if an action were judged by its consequences and not by the intention from which it arose. (1925/1965: 58, my translation)

Statements about translations by those who produce them – and more particularly, statements about the intentions behind the production – can be obtained from translators' prefaces, translators' blogs and think-aloud protocols, and also during retrospective interviews and professional development workshops. The statements may be

made some time after the wording of the translation is produced (prefaces, retrospective interviews, blogs) or at about the same time as production (concurrent think-aloud, comments made during group exercises in workshops). In this article I will be proposing a third kind of near-simultaneous commentary. Specifically, I will suggest making available to translators a vocabulary of one-word descriptors with which they can express, while they are translating, their beliefs about the source text and their intentions in writing a certain wording in the target language.

In using the vocabulary, translators would be engaging in self-analysis. They would be both observer and observed. They would be describing their own language production from a vantage point within their own minds while they are translating. In so doing, I will suggest, they would be able to reveal facets of translating that might well not come to light through the efforts of external observers.

No comparison will be made here between the proposed method and the familiar methods of concurrent think-aloud and retrospection. That is because the proposal is not specifically designed to be an alternative or a supplement to these methods (though it could perhaps be so used – see the section on uses of the vocabulary at the end of the article).

I have not trained a group of translators in the use of the proposed vocabulary, or even attempted to use it myself while translating. What I present is simply a suggestion, the value of which lies in giving me an occasion to bring out the potential role of translators themselves as investigators of their own translation process, and at the same time set out a certain conception of how a translator produces language.

Of the various text passages which I discuss in order to introduce sample items from the proposed vocabulary, all are authentic (though sometimes reworded to facilitate exposition) and all but one come from my own work as a translator. In keeping with the theme of the article, the thoughts I report about these texts are my own,[1] and they are presented in first-person, present-tense narrative style. For expository purposes, I sometimes divide myself into several personalities expressing alternative thoughts about a passage.

2. Ways of producing language

It is useful – as I hope to demonstrate here – to conceive of the activity of translating as a sequence of different ways of producing language, and I shall be illustrating some of these as I go over several examples of items in the proposed vocabulary.

Here is an initial example of how the vocabulary works. As I translate a medical text (a doctor's report to a colleague who had referred a patient), I mostly write in my

1. Some are thoughts which I emailed to myself while I was translating; others occurred to me while I was writing this article. The difference in timing does not matter since this article is not a report on any attempt to actually implement the proposed method.

own style, but occasionally I decide to imitate a doctor's style to add authenticity. In other words, I change the way in which I am producing language. When I make this switch, I enter the word VENTRILOQUISING in small caps in the text;[2] when I return to my own style, I enter END VENTRILOQUISING. VENTRILOQUISING is an item in the proposed vocabulary; it means "I am now attempting to imitate a style which is not my own".

The point in the text where I switch from my own medical style to what I believe to be a style used by doctors is the point where I *intend* to start such imitation. I may of course not succeed because I have failed to master "doctorese". An external observer (a reader or reviser of the translation, a researcher) may then see no "doctorese", yet I was still VENTRILOQUISING doctors' style because that was my intention. Or the opposite may happen: An external observer may identify the wording as authentic "doctorese", but I did not think of it that way; thus I was not VENTRILOQUISING, because I saw myself as writing in my own medical style, producing the words which I myself would use to talk about the medical matter in question.

In entering the word VENTRILOQUISING in the text, I am not providing a researcher with data revealing my current thoughts. After all, it is highly unlikely that a translator would ever spontaneously think of the word VENTRILOQUISING while translating. Rather I am engaging in second-level thought – metathought, if you like. The word VENTRILOQUISING represents my own analysis of what I am doing as I move through the text, an analysis that is channelled through an item selected from a pre-established set of words. Importantly, as noted in the previous paragraph, only I as the translator can know whether VENTRILOQUISING is going on; given the way it is defined – in terms of intention – no external observer can know.[3]

3. Stating intentions about adding

When external observers compare source and translation, they often find that the translator appears to have added material even though the addition is not a requirement of the target-language system. Suppose I am translating an in-house newsletter for employees of the ministry where I work. I find myself writing *there was a visitor*

2. This can be done very quickly using the AutoCorrect function of Word: typing "v-" would cause the word VENTRILOQUISING to appear on screen, with highlighting if so desired. In addition, a prompt could be made to appear on screen at intervals to remind the translator to enter items from the vocabulary.

3. It is true that with concurrent think-aloud, a translator might report an intention to imitate doctorese, but the translator would probably only think to mention this if the imitation seemed to resolve some sort of difficulty. In the method suggested here, translators would report VENTRILOQUISING *every time* they are aware of switching to it, whether the passage is causing a problem or not.

from headquarters this past week – a translation which conveys the meaning but is, like the rest of the text, dreadfully boring. So I decide to start producing language in a different way. To liven things up, I delete what I have just written and produce *everyone is wondering who this week's mystery man from headquarters was*.

An external observer comparing source and translation could see that I have added meaning, but such an observer could not tell whether I believed that the added meaning was compatible (or at any rate not incompatible) with the views of the source text writer. In other words, was the adding intended to be in the spirit of the source text? This is a matter of the translator's beliefs about the meaning of the source text, and intention in writing certain words in the target language. If I believe that the source has a certain meaning and if I intend my target-language wording to be in the spirit of the source, as in the above case, I enter the word FICTIONISING; if not (if I am engaging in deliberate deception about the meaning of ST, whether for laudable or nefarious purposes), I enter FABRICATING.[4]

Two questions may arise about additions: Was the addition intentional, and if it was, did the translator think that the added idea was implicit in the source? Suppose a passage of the source text (notes included with a CD recording) reads:

(1) *Prince Igor* is a well known Russian opera. Several famous Russian composers wrote operas on historical themes.

Here is the imagined thinking of two translators, both of whom are considering translating the first sentence as if it reads *Prince Igor is a well known Russian historical opera*:

Translator 1: Igor was a figure in Russian history, and the opera is about his doings, and it's an example of a historical opera, and the next sentence is about historical operas, so the notion "historical" is implicit in the source.

Translator 2: The writer doesn't want to downgrade Borodin's achievement by saying that Prince Igor is only well known within the subcategory of historical operas, but rather wants to say that this work is well known within the larger category of Russian opera as a whole. That's why there are two sentences, the first about *Prince Igor*, the second about historical operas. So the concept "historical" is not implied in the first sentence. If I translate here as if the first sentence contained the word *historical*, that will be me replacing the author's meaning with my own. Nevertheless, I'm going to do just that, because historical operas are extremely popular among opera lovers in my target culture these days, so I want to emphasise that word by having it appear in two successive sentences.

4. Readers should not make too much of these words. "Fabrication" and "fiction" often have very negative connotations in English, but these do not apply here. More generally, existing meanings and connotations of the words I have chosen for the vocabulary should be ignored. Their meaning is as defined here.

Using our pre-defined vocabulary, the first translator would enter the word EXPLICITATING to indicate that he thinks the notion "historical" is implicit in the source. The second translator would enter FABRICATING to indicate that he sees the notion "historical" as an outright addition, not implicit in the source, and not even in the spirit of the source. A third translator might enter no word from the vocabulary at all even though he had translated the sentence as if it contained *historical.* This absence of any word from the vocabulary would – ideally at least – indicate that the translator is unaware of engaging in any special kind of language production; that is, he has no intention to add. (Of course, another possibility is that the translator simply forgot to enter a word expressing his intentions.)

4. Stating intentions when uncertainty arises

Let us now look at some words that can be used to express intentions when the translator is experiencing uncertainty about the meaning of the source text.

If I am uncertain about which of two or more possible meanings a passage has, I may be able to write something which I believe allows both meanings; if I do this, I enter the word FUDGING. On the other hand, if I cannot see what the writer means at all, I may simply take a stab at the meaning; I then enter the word GUESSING. Note once again that in both cases, an external observer cannot tell what is happening (unless, in the case of GUESSING, I place question marks around the uncertain passage). Such an observer may not even notice any ambiguity in what I have written, but if I believe it has two meanings – if I intend ambiguity – then I am FUDGING. Conversely, the external observer may see an ambiguity in what I wrote even though I did not intend it; in that case, I was not FUDGING. A similar story can be told about GUESSING. GUESSING and FUDGING will be illustrated in the next section.

A further possibility when experiencing uncertainty about the meaning of the source is to engage in GLOSSING. The word GLOSSING means stating, in the target language, the values of the source-language linguistic units which compose the stretch of source text currently under consideration. It contrasts with the more usual method by which translators produce language, namely QUOTING. QUOTING means reporting to target-language readers more or less all of what the translator believes the source text means to its readers, and nothing more than that.[5] In entering the word QUOTING, the translator indicates that he intends to engage in metatextual activity, whereas the word GLOSSING signals metalinguistic activity.

Consider a French agricultural text about apple harvesting which refers to the problem of apples receiving *blessures mécaniques* ("mechanical injuries"). Below is the imagined thinking of four translators at that point in the text. The statements of the thoughts are intended to capture the logic of an argument; they should not be seen

5. Quoting is what many people would think of as "straight translating".

as wordings that might be produced if the translator were asked to think aloud. Each statement includes one or more expressions from our pre-defined vocabulary, and is followed by a brief commentary in parentheses.

Translator 1: I think it means the apples have mechanical injuries. This being my interpretation, I'm QUOTING the source by writing *[the problem of apples receiving] mechanical injuries*. (This translator sees nothing problematic about the English collocation *mechanical injuries* or its meaning, and does not pause for further reflection.)

Translator 2: I'm completely mystified about the meaning of *blessures mécaniques*, but I don't want to engage in GUESSING. So I'm GLOSSING by writing *[the problem of apples receiving] mechanical injuries*. (Here the translator does not have any interpretation yet and therefore simply writes definitions of the French words in English. In this GLOSS, the words *mechanical* and *injuries* state in English the dictionary meaning of the French words *mécaniques* and *blessures* respectively.)

Translator 3: I'm quite sure I know what is meant. I can see in my mind's eye apples that have been affected by this problem and I'm sure there's a special word for this in English but I can't for the moment think what that word is. So I'm GLOSSING by writing *[the problem of apples receiving] injuries of a mechanical rather than infectious nature*. (Here the translator has an interpretation, but it is strictly visual and therefore cannot be turned into a report in English of what the source means. The translator is not yet in a position to QUOTE and instead decides to GLOSS the French words by writing not one word for *mécaniques* but a sort of definition that spells out what is excluded by *mechanical*.)

Translator 4: I think it means there's mechanical damage. But I also think from context that they're talking about a particular type of damage, namely bruising, though it's true this notion would usually be expressed by *meurtrissures*. I'm going to QUOTE, reporting this more specific interpretation of the source by writing *[the problem of apples] getting bruised*. (Unlike in (1), the translator QUOTES only after considerable reflection.)

Notice how identical wordings are produced by translators 1 and 2. An external observer has no way of knowing what the translator intends (whether to QUOTE his interpretation of the source or to GLOSS the French words). If a translator were trained to use the vocabulary suggested here, 1 and 2 could be distinguished (QUOTING versus GLOSSING). However, no distinction would be made between 1 and 4 (both would be QUOTING) or between 2 and 3 (both would be GLOSSING). The proposed vocabulary does not distinguish every difference in thought process.

5. Stating intentions about linking words

Knowing the translator's intentions may be of interest with respect to links between sentences or between the parts of coordinate sentences. It frequently happens, in particular with poorly written texts and expert-to-expert texts, that the nature of the link is not at all clear in the source text. Consider this passage from an expert-to-expert

text written by a man's doctor to another doctor who works for the man's insurance company. The insurance company doctor has to decide, on the basis of a translation, whether the man is entitled to disability benefits as a result of being unable to work:

(2) The patient is taking cortisone injections for his lumbar pain. In principle I agree with the findings of the physiotherapist. Physiotherapy is more effective when pain symptoms are under control, and we have not really seen significant benefits with physiotherapy in this patient's case.

The translator does not have access to the findings of the physiotherapist, and the translator's client does not want the patient's doctor to be contacted for clarification. Below are some possible thoughts and subsequent actions concerning the word *and* in the third sentence in (2).

Translator 1: I have a couple of possibilities in mind for the connection between the two parts of this sentence, but I can't tell which of them is right. My target language won't allow the word for *and* here, so I'll write something ambiguous and leave it to the reader to figure out the connection. The easiest thing is to eliminate the connector and write two sentences, even though my target language has a rhetorical preference for overt connectors. I'm FUDGING.

Translator 2: I suspect the meaning here is concessive: Even though physiotherapy is more effective when pain is under control, we are still not seeing benefits from therapy. However, I'm not sure; this interpretation only works if the cortisone shots are controlling his pain, because then, the doctor says, you expect to see benefits from physiotherapy, benefits which have not in fact materialised. But we're not told whether the cortisone has been effective. The rhetorical norms of my target language don't allow the word for *and* here, so I'm going to use a concessive structure for now. Maybe things will be become clearer later. I'm GUESSING.

Translator 3: My target language allows the word for *and* here but I'm sure the meaning is concessive. What else could it be? So I'm QUOTING the source text – reporting to you my readers what I believe to be its meaning, but with specification of the connection between the two parts of the sentence, using the word for *although*.

Translator 4: I don't like FUDGING, so I'm not going to write two sentences with no connector. My target language does not allow the word for *and* when it is clearly not a matter of simple coordination. So I need to decide the connection between the two parts of this sentence. I've had this treatment myself (cortisone and physiotherapy) and my doctor mentioned something like this. What he told me was that physiotherapy is more effective when pain is under control, but the cortisone shots were not really working for me, and therefore I was not seeing benefits from physiotherapy. I'm going to translate here with a cause-effect structure, as if the source reads *therefore we have not really seen significant benefits*. However I don't think the notion "therefore" is implied in the source text. The meaning of the word *and* depends on whether this patient is benefiting from the cortisone shots, and nothing in the source text answers this question. So I'm writing something which I think is in the spirit of the source, based on my own experience. I'm FICTIONISING.

Translators 2 and 3 have both written concessive sentences, and translator 1 may be thinking of concession as one possible meaning. However, each is engaging in a different act: GUESSING, QUOTING and FUDGING respectively. Translator 4 is operating very differently: he is neither certain (like Translator 3, who QUOTES) nor uncertain (like Translators 1 and 2, who FUDGE and GUESS) about the meaning of the source; instead, he conveys his own meaning by FICTIONISING.

6. Some details about the proposed vocabulary

So far in this article, I have illustrated eight different language production acts: EXPLICITATING, FABRICATING, FICTIONISING, FUDGING, GUESSING, GLOSSING, QUOTING and VENTRILOQUISING. Each act is defined in terms of certain beliefs and intentions. Thus if I am FUDGING, then I believe that the meaning of the source is uncertain as between two (or more) possible interpretations, and I intend to write words in the target language which, I think, allow both interpretations. Here are some further details about the proposed vocabulary.

Completeness. The list of acts is not complete. Examples of further acts would be RECONSTRUCTING ("I'm correcting what I believe to be a defect in the source text" – perhaps a factual error) and SUMMARISING ("I'm deliberately not conveying all the ideas in the source, not even leaving them implicit"). The vocabulary has no pre-determined size; anyone who used it could add items to the ones I have given, or merge items.[6]

Structure. Some of the items in the vocabulary happen to be related, such as FICTIONISING and FABRICATING, but the list is not intended to have internal structure; there are no major categories or subcategories.

Simultaneity of acts. While some pairs of acts are mutually exclusive, others are not. One cannot be simultaneously QUOTING and SUMMARISING, since the definition of QUOTING excludes any intention to omit meaning which one believes to be present in the source. However, one can certainly QUOTE and VENTRILOQUISE at the same time.

Transformation of acts. Since the words in the vocabulary all refer to intention, it follows that the word appropriate to expressing intention about a passage of the translation can change at any moment. A phrasing which a minute ago constituted QUOTING can thus suddenly become an instance of SUMMARISING: if the translator has come to believe that the target-language phrasing conveys less than what the corresponding passage of the source means, but nonetheless decides to make no change, then the intention is now to SUMMARISE.

---

6. One might, for example, decide to treat EXPLICITATING as simply an aspect of QUOTING: if QUOTING is reporting all and only the meaning of the source as the translator sees it, then anything seen as implicit in the source, and made explicit by the translator, could be considered to count as QUOTING.

Unincluded acts. Language production acts that are transparent to an external observer are not included in the vocabulary. For example, at any moment I might switch the way I am producing language by pasting in a sentence from an earlier translation that has been found by Translation Memory. However the fact of pasting will be obvious to an observer who has recorded my translation session; its occurrence does not depend on beliefs and intentions, and therefore "pasting" is not a part of the proposed vocabulary. In the case of the "mystery man from headquarters" in the employee newsletter discussed above, it would be clear to an observer that the translator had engaged in adding, and therefore "adding" is not included in the vocabulary; however, it was not possible for such an observer to know whether the addition was a case of FABRICATING or FICTIONISING. Similarly, if a translator changed a wording in the draft translation, the fact of rewording would be obvious to an external observer, but it would not be possible for the observer to know whether the translator was REVISING or EDITING. REVISING means "I'm changing the translation because it fails to reflect something I now believe about the source text", whereas EDITING means "I'm changing the translation because of something I now believe about its linguistic features, style or structure". In other words, REVISING is motivated by the source text, EDITING by the draft translation. Even if an external observer knew through eye-tracking data that the translator had looked at the source text just before making the change, it would not follow that the change was an instance of REVISING rather than an instance of EDITING. The change might not be motivated by that glance at the source text at all. Which of the two acts is occurring will always depend on the translator's beliefs and intentions.

7. Uses of the vocabulary

I stated at the outset that the purpose of the vocabulary is to enable translators to self-analyse. In this way, they could make their own contribution to research. Having tagged several texts with items from the vocabulary, a translator could then write an article on a topic such as explicitation and use, as data, the EXPLICITATING tags which he/she had inserted in a set of translations. Such an article might be of considerable interest because the author could distinguish the tagged passages, where explicitation had been conscious during the translation process, from others, where the fact of explicitation only becomes evident after the fact, during analysis. It might also be of interest to compare a translator's self-analysis with an analysis by someone else of the same set of tagged texts.

The method was not specifically designed to furnish external observers/researchers with clues to mental processes, but it might nonetheless lend itself to such use. It might even prove to have advantages over existing methods: more systematic reporting of intentions than with think-aloud; consistency among subjects, since they would all use the same vocabulary; and a much smaller body of data to examine than with think-aloud protocols or retrospective interviews. On the other hand, subjects

would have to remember the names of the acts, and they would have to channel their thoughts through a set of pre-determined concepts, with the result that the external observer would not have access to the translator's personal way of thinking about a passage in the source or a wording in the translation. Despite these drawbacks, external observers might be able to use the self-analyses expressed through the proposed vocabulary. For example, they might ask several subjects to translate and tag a given text, and then look for differences among them in their recourse to FICTIONISING, FUDGING, etc., when uncertainty arises about the meaning of the source text.

Trainers too might be able to make use of the vocabulary. One of the great difficulties in helping new translators improve lies in determining why they wrote what they did. Often they cannot recall what they were thinking. If trainees were to tag texts with words like GUESSING or GLOSSING, the trainer might gain some insight into their intentions and then perhaps be able to diagnose certain problems in the trainee's approach.

Finally, the method of inserting items from a set vocabulary could be extended for the purpose of investigating matters other than intentions. For example, translators could express their views on the contents of the text they are translating (DISAGREE, INTERESTING), or what they are feeling as workers (BORED, TIRED). These aspects of translational experience could then be taken into consideration.

To conclude, the proposal sketched here has potential to allow translators to self-analyse by stating their beliefs about the meaning of the source text as well as their intentions in writing a particular wording in the target language, and to do so while they are in the process of writing. In presenting the proposal, I have also introduced two ideas which I think are novel. First, the notion of switching between different types of language production as the translator moves through a text is a way of thinking about translating which, I believe, has not been suggested before. Second, the composing work of the translator has been described here in a way that I think is more realistic than the traditional pseudo-objective manner. What I have called QUOTING would usually be described as the translator conveying "all and only" the meaning of the source text. But "all and only" according to whom? A heavenly panel of omniscient revisers and translation teachers? It is important to see that when translating is taking place, the only person whose beliefs and intentions are relevant is the translator.

Reference

Bally, C. 1925/1965. *Le Langage et la vie*, 3rd edition. Geneva: Droz.

Metonymic language use as a student translation problem

Towards a controlled psycholinguistic investigation

Sonia Vandepitte*, # and Robert J. Hartsuiker#
*University College Ghent / #Ghent University, Belgium

The concept of translation problem plays an essential role in translation process and product studies and encompasses a wide variety of kinds of problems, including those of a linguistic nature. While Translation Studies seems to agree that a linguistic competence module is part of a translator's required competences, it is not clear to what extent cross-linguistic differences actually pose problems to most translation students and therefore need a place in the training curriculum. Our study looks into the extent to which the translation of a linguistic construction that differs cross-linguistically raises translation process problems. Translators who are not aware of the restrictions (or translation tools that have not integrated the cross-linguistic differences) may, indeed, produce a translation that does not conform to the target language idiom. And translators who are aware of the restrictions will need to look for alternative formulations that are suited to the translation situation, and decide on the best translation from a range of possibilities. This chapter will discuss the methods applied in an investigation of a novel kind, which combines approaches from linguistics, psycholinguistics and Translation Studies.

Keywords: translation problem, metonymy, agent prototypicality, English-Dutch, response latencies

1. Translation problems

In translation, various problems arise and need to be solved. Such translation problems are a central issue not only in translation process studies but also in translation product studies (Nord 1991; Chesterman 1997; Göpferich 2000, 2008; Englund Dimitrova 2005; Gambier 2008). They may belong to a wide range of events or states, witness Thome's (2004) classification of translation problems. Hence, for each language combination and translation direction, repertoires can be set up of translation

problems of various types – "lexical, syntactic, pragmatic, text-linguistic, culture-specific, creativity-demanding and comprehensibility-related problems" (Göpferich 2009). They constitute so-called "Rich Points" according to the PACTE group (e.g. PACTE 2009). While translation scholars agree that among a translator's required competences, the linguistic ones constitute one of six basic pillars (PACTE 2000; Göpferich 2008: 155), it is not clear to what extent cross-linguistic differences actually pose problems to most beginning translation students and therefore need a place in the training curriculum. A potential instance of such a rich point is where languages differ as to the distribution pattern of metonymic and non-metonymic use of items in a certain construction. A case in point is the occurrence of a non-prototypical agent[1] with a prototypical agent requiring predicate (abbreviated here as NPAPARP): this is the combination of an abstract or non-human noun that typically causes or helps cause the action expressed by the predicate, with a predicate that usually requires a human agent. When such an NPAPARP occurs in a source text language that allows for it and has to be translated into a target language that imposes restrictions on such constructions, there is a potential translation problem. When non-metonymic, an English source text structure can be copied as a Dutch target structure (1); however, its metonymic occurrence does NOT allow such copying, as in (2) and (3):

(1) English source: This man might hurt the entire population of the City of London.
Dutch translation: Deze man zou de hele bevolking van de Londense City kunnen kwetsen.
(Gloss translation: This man would the entire population of the City of London be capable of hurting.)

(2) English source: Money buys action and influence.
Dutch translation 1: ? Geld koopt actie en invloed.
(Gloss translation 1: ? Money buys action and influence.)
Dutch translation 2: Met geld koop je actie en invloed.
(Gloss translation 1: With money buy you action and influence.)

(3) English source: An EU savings-tax directive […] might hurt the City of London.
Dutch translation 1:? Een EU-richtlijn met betrekking tot belasting op spaartegoeden zou de Londense City kwetsen.
(Gloss translation 1:? An EU savings-tax directive […] might hurt the City of London.)

1. Agent is used in its broad sense to include both the person performing the action expressed by the predicate and the object that helps perform the action, sometimes also referred to as the instrument. An agent ("this man") is not necessarily the subject of a sentence ("The City of London"), as in "The City of London was hurt by this man".

> Dutch translation 2: De Londense City zou lijden onder een EU-richtlijn met betrekking tot belasting op spaartegoeden […]
> (Gloss translation 2: The City of London would suffer from an EU directive with respect to tax on savings.)

Although product Translation Studies have discussed translation problems by focussing on quality issues and process Translation Studies by discovering pauses, hesitations, problem-solving cues by means of think-aloud protocols, questionnaires, logging and eye-tracking, investigations that deal with a particular syntactic-semantic structure (like the NPAPARP above) are still missing. In addition, they are not only absent in Translation Studies but also in psycholinguistics, so that the basic question still remains as to whether translation behaviour confirms the occurrence of a "real" translation problem. Consequently, the teacher who needs to determine whether or not to include training in contrastive syntactic-semantic structures into a translation training curriculum is still at a loss. In other words, a controlled psycholinguistic investigation is in order. This chapter reports on an ongoing experimental study into the translation of such metonymic use of language and inquires into the extent to which English metonymic language use of non-prototypical agents in a prototypical agent requiring environment results in higher translation processing costs (i.e. longer translation times) than the prototypical agents in a similar position. Additionally, this chapter compares findings among beginning translation students with those among advanced translation students.

With such an aim, this study is to be situated among studies that inquire into translation interference, which started to appear in the second half of the twentieth century and have mainly relied on *product* findings in which observers recognize L2 features in L1 translations. Initially, anecdotal evidence was brought forward (Nida 1964; Catford 1965), but, later, studies looked into general trends (Toury's law of interference 1995) and translationese features in large translation corpora (e.g. Mauranen and Kujamäki 2004). It is only in the last two decades, however, that investigation into the phenomenon has started in psycholinguistics, mainly looking into the influence of a bilingual's first language (L1) on the processing of a second language (L2) or vice versa. Such studies have demonstrated that processes such as visual word recognition in reading (e.g. Dijkstra, Timmermans and Schriefers 2000; Van Assche, Duyck, Hartsuiker and Diependaele 2009), auditory word recognition in listening (e.g. Weber and Cutler 2004), and syntactic processing in speaking (e.g. Hartsuiker, Pickering and Veltkamp 2004) are affected by cross-linguistic overlap in form and/or meaning. Studies on reading, for example, reveal that cognates (EN *film* – NL *film*) can be recognized more quickly than matched non-cognate control words (e.g. Caramazza and Brones 1979; Duyck, Van Assche, Drieghe and Hartsuiker 2007). This is also the case when the participants read in their L1 (van Hell 2002) and even when the cognates are embedded in sentences (Van Assche, Duyck, Hartsuiker and Diependaele 2009). In many of these studies, reading times are determined by measuring the time to execute a word/non-word decision (lexical decision); but an increasing number of studies use

the (arguably much more ecologically valid) method of registering eye-movements during reading (Duyck, Van Assche, Drieghe and Hartsuiker 2007; Libben and Titone 2009; Van Assche, Duyck, Hartsuiker and Diependaele 2009). Similarly, studies on listening have demonstrated that cross-lingual overlap in the sound of words influences which objects bilinguals look at. Thus, when Weber and Cutler (2004) presented Dutch students with English as an L2 with the auditory word "desk", they more often looked at a picture of a lid ("deksel" in Dutch) than at an unrelated object. Importantly for the present purposes, studies on language production show that speakers tend to persevere in their choice of syntactic structures, *even across languages*. Hartsuiker, Pickering and Veltkamp (2004) demonstrated that when Spanish-English bilinguals first hear a Spanish sentence in the passive voice and then describe a picture of a scene in English, they are more likely to use an English passive than when the initial sentence in Spanish was in the active voice. Similar cross-linguistic "structural priming" effects were observed with the dative alternation (Schoonbaert, Hartsuiker and Pickering 2007) and with noun modifiers, as in *the shark that is blue* vs. *the blue shark* (Bernolet, Hartsuiker and Pickering 2007). Anecdotal evidence from such experiments suggests that structural priming not only biases speakers to choose a given, syntactically acceptable option in the target language; it can even seduce speakers to (unacceptably) transfer a particular structure. Thus, in one of Bernolet's experiments, bilingual participants were sometimes observed to produce phrases like "the cowboy his hat" in English instead of "the cowboy's hat", presumably because they had just been primed with the Dutch construction "de cowboy zijn hoed".

A small number of these studies in psycholinguistics (Christoffels, de Groot and Kroll 2006; Ruiz, Paredes, Macizo and Bajo 2008; van Hell and de Groot 2008) have used both word and/or translation tasks (and, in some cases, tested professional translators or interpreters). The current study also has its participants perform a word translation task (pre-test) and a sentence translation task (test), with the aim of detecting whether features of the source language affect the time course and form of production in the target language. Although word and sentence translation tasks may be controversial in Translation Studies for lack of "ecological validity" (on its importance, see e.g. Shlesinger 1998: 489–490), they have been preferred to text translation tasks, which would not allow for a similar rigid interpretation of the findings because of the complexity of the contextual influence on the structure under investigation.

2. Npaparps

Npaparps turn out to be constructions whose acceptability differs cross-linguistically and thus exemplifies how conditions for relationships among sentence parts are determined within each language. In the second translation of (2), for instance, the noun phrase *je* is determined by its relation to the other sentence parts, which have assigned it its specific place and form (the grammatical function of *je* is subject) and its meaning in the sentence or its thematic role (*je* is agent). The relationship between the

subject and the agent in all languages is known to be a privileged one. Dowty's (1991) Argument Selection Principle states that predicates which have both a subject and an object will have a subject that fulfils most Prototypical Agent features:

a. volition
b. sentience and/or perception
c. causation
d. movement and
e. independent existence.

The more prototypical agent features a noun phrase has, the larger the likelihood that it will take subject position. This means that there is a hierarchy, in which *human* nouns (e.g. *je* in (2)) with their typical agent characteristics will be at the top of the list of privileged candidates. It also implies that less prototypical agents (e.g. *money* in (1)) will occur in subject position less frequently (like *geld* in (1)). However, languages seem to have different degrees of tolerance with respect to this. English, for example, is fairly flexible and easily accepts non-human nouns that fulfil the Proto-Agent features to a lesser degree or not at all, as is shown in (2) and (3). Such structures are marked, however, in Dutch, as already pointed out by Vandepitte (2007): non-prototypical agent subjects are less acceptable in Dutch and, mostly, speakers will opt for different constructions as in the second translations in ((2) and (3)).[2] Such metonymic language use is also marked in other languages: in German (Hawkins 1986: 175; Bahns 1993; Göpferich 2004; Baumgarten 2005), Spanish (Slabakova and Montrul 2002) and Asian languages (Master 1991; Lee 2001), too, non-prototypical agent subjects are less acceptable. In other words, languages vary with respect to the restrictions that apply to the use of non-prototypical agent nouns in subject position and the way in which such nouns are integrated into the sentence structure. Language restrictions may also show a gradation conflict: in certain contexts, literal translation is acceptable as metonymic use.

This chapter presents the results of an ongoing investigation into whether translations as in (2) and (3) are indeed problematic for an English into Dutch translator who will need to obey target language norms and who will not be able to translate such non-prototypical agent subjects as similar non-prototypical agent subjects in the target text, or in other words, whether the translator faces a linguistically observable semantic-stylistic target language related translation problem (Thome's classification 2004).

2. The NPAPARPs in the examples need to be distinguished from obvious personification examples such as "Now I assume that Mother Nature will age me along with everyone else" (BNC, ACP 1327), in which the capitalization of the subject (and possibly an additional noun denoting a person like Mother) clearly indicates that the subject ought to be seen as a prototypical agent with all its characteristics.

3. Research questions and hypotheses

Although the problem of NPAPARPS may be recognized by English language teachers and translators, it has not yet been confirmed by translator behaviour. Until now, no process study has shown that student translators really have problems with these structures. In psycholinguistic terms, we can reformulate this behavioural question as follows: Is a student's translation affected by priming, a question that is operationalized in terms of the reaction times (*latencies*) in which translations are produced? Or does priming by English non-prototypical agents in a prototypical agent requiring environment really take place? And since training may play an important role in a student's translation competence, we can also ask the question whether priming by English non-prototypical agents in a prototypical agent requiring environment takes place with both untrained and trained students (Q1 and Q2 below):

Q1: Does the occurrence of non-prototypical agents in a prototypical agent requiring environment in English prime translation into Dutch with untrained students?
Q2: Does the occurrence of non-prototypical agents in a prototypical agent requiring environment in English prime translation into Dutch with trained students?

Non-prototypical agents often are abstract nouns, and psycholinguistic investigation already informs us that there is a tendency for concrete items to prime a linguistic activity across languages more strongly than abstract items (Schoonbaert, Duyck, Brysbaert and Hartsuiker 2009: 573). More important for the present investigation, such stronger priming by concrete words also occurs in isolated word translation tasks (van Hell and de Groot 2008). In view of the different degrees of abstraction between the two types of agents in our own study, we would therefore hypothesize that latency will be longer for the translation of isolated non-prototypical agents than for the translation of isolated prototypical agents. Since van Hell and de Groot show how the context of the item influences the degree of priming but not its nature – the difference between the priming effects by concrete and abstract items being stronger when tasks are performed on items that can*not* easily be predicted from the rest of the sentence (as in *Many researchers are still looking for* ***the truth***) than when tasks are performed on words that can be easily predicted (*After some hesitation my father told* ***the truth***) – we may assume that our hypothesis is also applicable to the translation of (non-)prototypical agents when they co-occur with a prototypical agent requiring predicate, even though the context of the agents in our investigation appears *after* the agents and will, therefore, *not* have any significant influence on their priming capacity. Summarizing, we hypothesize that the average latency per person will be longer for the translation of sentences with non-prototypical agents than for the translation of sentences with prototypical agents.

If we want to answer the main question whether NPAPARPS actually equally constitute a translation problem for untrained and trained students, this study will investigate latency data of both groups. In the ideal didactic situation, we expect the problem to be smaller for the trained students: trained students are expected to show

shorter latency than untrained students as well as a larger variety of alternative structures rather than the literal translation.

4. Experimental design

In order to solve the questions, we need to be certain that any differences in latency are differences due to the NPAPARP constructions themselves rather than any potential differences between the prototypical and non-prototypical agents. That is why a controlled pre-test was set up to determine the extent to which differences can be observed between the processes in which *isolated non-prototypical* agent nouns are translated and those in which *isolated prototypical* agent nouns are translated (see also Section 5.2 below).

The main test itself required participants to translate 76 sentences. It has a between-subject two-by-two factorial design (non-prototypical / prototypical agent; untrained / trained) in which each participant saw each stimulus in each condition in a different random order to offset any influences that participants' learning from preceding examples may exert on the translation processes.

5. Test

5.1 Aim and expectations

The aim of the test is to identify the speed with which NPAPARPS and PAPARPS are translated by both untrained and trained students. In such a study, participants are given non-target language representations, and the aim is to establish the reaction times in which they start their translations. In order to exclude the possibility that a particular translation construction requires a certain reaction time whatever the input for translation is, we will also relate reaction times to translated constructions. Test measures will therefore be onset time and translation construction type and we will relate those to the variable of student type (untrained vs. trained).

We expect priming to take place in both groups of students: they will all need more time to translate the NPAPARP constructions than the PAPARP constructions (the untrained ones even more so than the trained ones). We also expect both groups of students to need more time for the production of Dutch NPAPARPS than for Dutch PAPARPS.

5.2 Materials

Although, perhaps ideally, materials should be derived from authentic professional translation assignments or from translation corpora (cf. Vandepitte's pilot corpus

2007), they are so divergent that they cannot possibly fulfil conditions necessary for experimental requirements which aim at excluding as many variable factors as possible. Psycholinguistic studies have shown that important factors relevant to the speed of visual word recognition are word frequency (Whaley 1978) and word length (McClelland and Rumelhart 1981).

It was therefore decided to use a paired set of 38 NPAPARP constructions (4) and 38 PAPARP constructions (5) with very specific nouns (for the whole set, see Appendix E):

(4) Music took him around the world.

(5) Mother took him around the world.

The *subject nouns* in this test and their translations had first been investigated as to their suitability in a pre-test, which cannot be fully described here. Noteworthy is that both 38 non-prototypical agents (critical words) and 38 prototypical agents (control words) were taken from the COBUILD corpus in the Celex database (Baayen, Piepenbrock and van Rijn 1993). We used Duyck, Desmet, Verbeke and Brysbaert's Wordgen tool (2004), which yields summated bigram frequency figures for all words, which are the sums of the frequencies of all sequences of two letters in a word in a particular language (Rumelhart and McClelland 1985). The critical words matched the control words in terms of the following characteristics: they are equally easy to translate into Dutch (they are all near-cognates with Dutch words, e.g. *father – vader*, with the exclusion of full cognates, e.g. *film – film*), they have a similar potential for ambiguity and polysemy (most of them are not homonymous with verbs or adjectives; cf. Tokowitz in Sunderman and Schwarz (2008: 529), who found that polysemy may slow down the translation process) and they are all countable and in the singular form requiring a predicate in the singular form (which excludes a word like *police* from the investigation). From those words, the ones with an American spelling were excluded (*center, color*), since the student participants are taught British English. More importantly, however, their word frequency is very high (they are all among the first 5% highest frequency ranks), their graphemic word length is similar (ranging between five to eight letters), their phonological word length is similar (ranging between one to three syllables) and their summated bigram frequencies are similar.[3] Independent sample t-tests over critical and control items showed no differences in word length, word frequency or summated bigram frequency (all $t < 1$, all $p > .05$). Similarly, independent sample t-tests over critical words and their near cognate translations in Dutch showed no differences in word length, word frequency or summated bigram frequencies (all $t < 1$, $p > .05$). Most prototypical words are human and individual (*brother, daughter, swimmer,* etc.), while the non-prototypical agents are abstract

3. It was assumed that the summated bigram frequencies for the British English words that have an equal spelling in both varieties of English are fairly similar to the ones for American English.

(*project, music, summer,* etc.), none of them denoting companies or other groups of people, nor any words that refer to quantities (*amount, series,* etc.), since semantically they are not the (non-prototypical) agents themselves (it is the entities they quantify that are the agents). The pre-test showed that the mean translation onset times did *not* differ at all: both the group of prototypical agents (human) and that of non-prototypical agents (abstract) had a mean of 1.198 s. In other words, if the translations of sentences containing these words do show different onset times, it means that not the words themselves but the contexts around them – and we assume especially the factor of agent-prototypicality – play a role. All words are presented in appendix B, and their lengths, frequencies and summated bigrams in appendices D1 and D2.

The critical *sentences* themselves (English NPAPARPS) were construed in such a way that they all had similar information distribution patterns and that the determiners and premodifiers of the agents in a sentence pair were similar (*Mother/Music, His career/His sister, Most of the tourists/Most of the action,* etc.; with two exceptions *plague/criminal* and *essence/speaker*). Particular attention was also paid to the type of predicate (predicates requiring plural subjects as in *the prayer flags and prayer wheels dot the landscape* were excluded) and the type of register in which they usually appear (NPAPARPS are particularly frequent in headlines such as *Climate change favours coffee pest*).

Each participant gets his/her own coded file to translate with the same mixture of critical and control words but presented in random order. Any effects from getting better at translating words towards the end of the task or from translating two different sets of agents (prototypical versus non-prototypical ones) can therefore be excluded.

5.3 Participants

Sixty-one students from University College Ghent Translation Studies who had participated in the pre-test were invited by means of mailings via the learning platform and oral messages in classes to take part in the test. Mailings invited volunteering students to participate twice, once before the Easter holidays (pre-test) and once after the summer exams (main test). Forty students turned up voluntarily but only 24 students produced results that were fully comparable: 10 students were untrained (first-year bachelor of applied languages students) and 14 trained (master of translation students).

Ghent lies in Flanders, i.e. the Dutch-speaking region of Belgium, and the students were all native speakers of the Dutch target language. After the main test they were given a cinema ticket as remuneration, and they received personal feedback on their performance in terms of speed and adequacy as compared to their peers. They took the test in what was a natural student translation environment for the trained students, namely the language laboratory class at a PC that had Morae software; for many untrained students, however, this was the first or second exercise in such a class.

Explanations and instructions were given to students in Dutch. They were told they would be given a short language history questionnaire inquiring into the students' native tongue, the number of years in which they have studied English, at which age they started to learn English, how much time they have spent in an English-speaking country, which other languages they know, what dialect or language they use at home, and how they assess their reading, writing, speaking and listening competences in English and translation from English into Dutch, their PC typing experience, speed and blindness, the keyboard and their vision (Appendix A). The questionnaire also gave them some space to comment on the experiment. Further, the students were informed that they would perform a half-an-hour translation task in which they would see words (pre-test) and sentences (main test) that they should translate out of context. Students were given various opportunities to participate in the test, most of them during lunch hour. Anonymity was ensured and the only aim revealed was improvement of translation classes.

Immediately before the experimental sessions, students signed the consent form and filled in the questionnaires. Then students followed the instructions of their handouts in order to start the Morae recording and their translation task in a Word document. Both files were saved by the students themselves onto USB sticks provided by the experimenter. To start the sentence translation task itself in the Word document, participants were asked to read the sentence they saw, to go to the next line by means of the "End" and "Enter" buttons and to type in the translation of the sentence as quickly as possible. Then, tapping the "PgDn" key twice immediately brought them to the following sentence which they again had to translate, and so on. They were specifically asked to remain focussed on the task, keep their attention constant and not to pause, as if they were translating for a client who was waiting to take these translated sentences along immediately after they had finished them.[4] They were asked to type the *first* translation that they found: the most spontaneous translation was the only one that was needed. They were instructed *not* to go back to a preceding word or to press any other keys (incl. the space bar) than the letters and "PgDn", "End" and "Enter". At the start, students were given five sentences typed in blue to practise typing, translating and passing to the next sentence, which did not reappear in the rest of the test.

5.4 Analysis and scoring

All participants' Morae recordings were analysed: the latency of each translation, whether adequate or inadequate, was measured by subtracting the time at which the second "PgDn" key was touched from the time at which the first letter of the translated sentence was typed. If students did not succeed at touching the "PgDn" key twice

4. Although this "translation" task is fairly infrequent, it regrettably does happen that a client asks a translator to provide the translation of materials without delivering any contextual information, whether intratextual or extratextual (Abdallah 2010: 20).

followed by "End" and "Enter", the translation onset time of the word that followed was considered not to be comparable to the others and therefore not incorporated in the data.

Translations were also scored in terms of the presence of a prototypical agent (6), a non-prototypical agent (7), no agent (8) and no explicit agent other than the non-prototypical agent (9):

(6) English source: His career focussed on economics.
Dutch translation: In zijn carrière legt hij zichzelf vooral toe op economie.
(Gloss translation: In his career applies he himself especially to economics.)

(7) English source: His career focussed on economics.
Dutch translation: Zijn carrière focuste op economie.
(Gloss translation: His career focussed on economics.)

(8) English source: The best practice combined two approaches.
Dutch translation: De beste praktijk was de combinatie van twee aanpakken.
(Gloss translation: The best practice was the combination of two approaches.)

(9) English source: That period put much emphasis on metaphors.
Dutch translation: In die periode werd er veel nadruk gelegd op metaforen.
(Gloss translation: In that period was there much emphasis put on metaphors.)

5.5 Results

Of the 1766 translations generated in the experiment, 918 (52.0%) had a prototypical agent, 696 (39.4%) had a non-prototypical agent, 66 (3.7%) had no agent, 3 (0.2%) had no explicit agent, and 83 (4.7%) had not been translated. Table 1 shows the distribution of the translations as a function of group and condition (with "no explicit agent", which had few observations, collapsed with "no agent").

Table 1. Percentage of translations in each category as a function of group and source language agent prototypicality

	N	P	NA	No translation
Untrained group				
Non-prototypical ST agent	83.3	8.4	4.4	3.9
Prototypical ST agent	1.3	90.7	0.8	7.2
Trained group				
Non-prototypical ST agent	72.7	15.2	8.0	4.0
Prototypical ST agent	1.8	92.3	1.8	4.1

Note. N = non-prototypical TT agent, P = prototypical TT agent, NA = no TT agent.

Table 1 shows that both groups had a very strong tendency to produce translations with a non-prototypical agent in the condition where the source-language utterance also had a non-prototypical agent. However, this tendency was more pronounced in the untrained group (83.3%) than in the trained group (72.7%). Responses without any (explicit) agent were observed too infrequently to allow for an appropriate statistical analysis, but at least descriptively, these occurred more often in the trained group's translation of non-prototypical agent than the untrained group's translations. For each condition, a Chi-square test was conducted on the distribution of non-prototypical and prototypical responses in each group. In the condition with non-prototypical source text agents, the untrained group produced more translations with non-prototypical agents than the trained group ($\chi^2(1) = 10.4$, $p < .001$). In contrast, in the condition with prototypical agents, both groups produced almost exclusively responses with prototypical agents and there was no difference between the groups ($\chi^2(1) = .047$, $p = .83$). Summarizing, both groups show a very strong tendency to translate *constructions with a non-prototypical agent with a construction that similarly has* a non-prototypical agent. But this tendency is weaker for the more advanced students, who more often appear to search for different solutions to the translation problem (such as sentences with a prototypical agent or agentless sentences).

We further analysed the translation latencies. The latency analysis was restricted to translations with either a non-prototypical or a prototypical agent (excluding 152 responses). A further 176 latencies were not registered appropriately. Finally, of the 1438 remaining observations, 28 (1.9%) were excluded as outliers, defined as being three standard deviations or more longer than the grand mean. In the remaining 1410 observations, the grand mean was 4.93 s (SD = 2.11). Table 2 lists translation latencies as a function of group and condition.

An ANOVA with group (trained vs. untrained) as a between-subject variable and agent prototypicality of the source sentence (prototypical vs. non-prototypical) as a within-subject variable showed that sentences with non-prototypical agents were translated significantly more slowly than sentences with prototypical agents, $F(1, 21) = 27.7$, $p < .001$. There was no effect of training, $F < 1$, and training did not interact with agent prototypicality, $F < 1$.

Table 2. Translation latency

	Translation time	
	Mean	**Standard deviation**
Untrained group		
Non-prototypical agent	5.22	.94
Prototypical agent	4.59	.66
Trained group		
Non-prototypical agent	5.33	1.39
Prototypical agent	4.58	1.01

Finally, (at least) descriptively, translations having prototypical agents were produced much more quickly when the source sentence also had a prototypical agent (M = 4.59 s, SD = 1.84 s, N = 709) than when they had non-prototypical agents (M = 6.15, SD = 2.86, N = 96). Unfortunately, there were too few responses in the latter category to allow for statistical analysis of this difference.

6. Conclusion

Summarizing, our findings show that the occurrence of non-prototypical agents with a prototypical agent requiring environment in English does prime translation into Dutch with both untrained and trained students: not only do students need more time to translate them than they need to translate PAPARPs, they also produce a prototypical agent construction more slowly if the source sentence is an NPAPARP than if the source has a prototypical agent. In other words, the NPAPARP constitutes a clear problem both for untrained and trained students, so we can conclude that the translation of this linguistic construction that differs cross-linguistically constitutes a translation process problem. In addition, the high percentage of trained students translating literally (more than 70%) and the relatively small difference that we find between untrained and trained students point at scope for improving translation training.

7. Methodological discussion

Let us now take a closer look at the methods employed in this investigation and reflect on potential alternative ways.

7.1 Pre-test

In the first place, the number of items in the pre-test (40 + 40) should perhaps ideally have been larger. Indeed, our results included a fairly high percentage of outliers (> 2%), which yielded a data set with more variability than we had hoped for: latencies that were more than 2.5 standard deviations above or below the participants' mean latency were not considered as outliers nor eliminated from our analyses (in contrast, for instance, with the trimming procedures that were applied in Sunderman and Schwartz (2008)), and, at this stage, it is not clear where to draw the line between acceptable and unacceptable data loss. However, the number of language items that fulfilled conditions as specified in Section 5.2 did not turn out to be limitless, and finding a set of critical items that were matched against both a set of control words and a set of potential translations proved to be a very delicate exercise. For the main test, however, we excluded the two items that turned out to be most difficult to translate for many students and their controls: for both *clerk* and *revolt* only 37 students (67%

of all participants) were able to give a translation within three seconds. With their exclusion, the t-values of the two sets of Dutch critical and control words are not significantly influenced. However, the two sets of English critical and control words are less matched for frequency, even though the t-value is smaller than 1. This is, in fact, a remarkable result. If students have difficulty translating these two words, it is because they are not familiar enough with the words to make sense of them without meaningful context, and both words have turned out to be equally unfamiliar to all 55 students. Yet, it turns out that their frequencies are different enough to affect the t-value.

This brings us to the methodological problem of the corpus chosen for the investigation and its frequency data and the participants in the investigation. The COBUILD database was produced more than fifteen years ago and contains American English produced by native speakers. The participants in the test, however, are much younger than those adults, have Dutch as their native language and are in the process of acquiring proficiency level of English. This problem has already been pointed out by Duyck, Van Assche, Drieghe and Hartsuiker (2007): they claim that the application of existing models of word recognition cannot be extended to unbalanced bilingualism, and Brysbaert's research team (Keuleers, Brysbaert and New, under revision) is taking up this problem: they are now drawing frequency values from a large corpus of television subtitles and claim that these values will predict better lexical decision times than the Brown corpus (Kucera and Francis 1967) or Celex.

7.2 Test

Although other factors than student type and agent type may play a role in the diversity of latencies established (general individual reaction speed could explain differences between students of the same type, individual student bigram frequencies, or even the place of the initial letter of the word and in particular the place of its key on the keyboard), the combined analysis of target time and construction choice has allowed us to conclude that the NPAPARP constitutes a translation problem and is therefore a valuable topic to be included in translation training. We have not yet answered the question, though, how this matter should be incorporated ideally. We do not know whether, for instance, setting up digital exercises similar to the experiment that also provide adequate and acceptable translations is a solution that allows for flexible self-study or whether it would interfere with the aim of translation training to have students focus on texts and derive meanings from texts. Nor can we decide that the small difference that we observed between trained and untrained students is due to training: since our trained students have been selected in the course of their student career, while the untrained ones have not, this selection may constitute a bias in our sample of the population. In addition, the data of the language history questionnaire still need to be processed in order to set up an average profile of the group and consider other participant factors that may play a role in the training issue. Finally, we do not know whether the NPAPARP construction constitutes a problem for translation

with other language pairs. In fact, our present findings concerning an individual's language skills may well be influenced by the fact that English is a *socially* dominant language. Pressure from English in Flanders is very large: among students, it may well have a higher prestige even than that of Dutch. In order to generalize our results, similar tests should be undertaken with a less socially dominant language such as German or French.

It must also be admitted that the ecological validity of this sentence translation task is a controversial issue in Translation Studies: although this test does provide some contextual information, a client asking a translator to translate isolated sentences is a rare one indeed. Is it at all possible to apply our experimental findings to a naturalistic student translation setting? Will students have similar difficulties translating the NPAPARPS in larger texts?

And yet, we consider this task as an essential step in building our knowledge on translation. With our findings we know that many NPAPARPS translations show similar latencies. A study focussing on a small number of longer individual translation assignments will never be able to reveal that type of information. Ours is a bottom-up approach, which assumes that studying the translation process of a word in isolation yields information relevant for the study of the translation process of a sentence in isolation, which in its turn yields information for the study of the translation process of a sentence in a specific textual context. A future within-subject design that is applied to our data will be able to compare latencies of an individual translator for isolated word and isolated sentence translations and test the first part of our assumption.

If that yields positive results, we can look into methods to investigate the translation of NPAPARPS in larger texts and take the next experimental step. This will raise challenging questions such as: Will it be possible in the future to test context-embedded translation processes? Could a text be created that contains a subset of the experimental sentences with primes and controls in similar and different contexts? Is it still possible to call this type of priming, in which there is a combination of semantic features with a specific structure, structural priming? What are the consequences for psycholinguistic bilingualism models? Alternatively, we could theoretically build huge corpora of translation time recordings and look for NPAPARP translation latencies, another long and difficult research journey.

Before we take this next step however, we first aim at finding out whether the artificiality of the translation assignment made participants conscious of any problems associated with the non-prototypical agents. To answer this question, an oral post-test survey is now under way with questions asking for general impressions about the translation, any particularly difficult matters and for clarification of any special translations or time delays. Preliminary findings already reveal that students realized that their translations were not very good and considered them the best compromise possible: instead of looking for a better solution in the target language, they decided not to lose time and hoped that timeliness combined with the adequacy of the translation would compensate for the lower degree of acceptability within the target language. Our investigation will therefore in future also take a closer look at quality issues and

investigate to what extent translation onset time is related to adequacy/acceptability. Indeed, the longer latency for NPAPARP translations than for PAPARP translations is not the only characteristic that is indicative of priming. Apart from latency, there is another factor in terms of which our research questions can be operationalized: quality. Our measure was translation *onset* time, but a closer investigation of participants' deletion of letters and words and of the translation times of the *complete sentences* and their relation to translation quality will also give us an insight into the self-monitoring process and how it may influence the quality/time ratios of translations.

Completely different questions also arise: Is there a relation between the experimental setting with two types of students and natural professional translation? Are our investigation data comparable with corpus findings? Answers to these and the previous questions will not only be useful for translation didactics but also for psycholinguistic modelling of bilingualism. They will offer the potential to extend psycholinguistic research by looking into priming of a construction in translation that is not only constrained by syntactic features, but also by semantic ones. The data collected will also reveal information about the translator's self-monitoring processes, i.e. another stage/activity of the complex process that translation consists of. In other words, the data will also help us understand the cognitive mechanisms that translators use to deal with conflicting forces on syntactic choice and contribute to our insight into "how linguistic context influences bilingual word recognition" (van Hell and de Groot 2008: 432).

References

Abdallah, K. 2010. "Translators' agency in production networks." In *Translators' Agency. Tampere Studies in Language, Translation and Culture*, T. Kinnunen and K. Koskinen (eds), 11–46. Tampere: Tampere University Press.

Baayen, R. H., Piepenbrock, R. and van Rijn, H. 1993. The CELEX Lexical Database (Release 1) [CDROM]. Philadelphia: University of Pennsylvania Linguistic Data Consortium.

Bahns, J. 1993. "Lexical collocations: a contrastive view." *ELT Journal* 47 (1): 56–63.

Baumgarten, N. 2005. "The Secret Agent: Film dubbing and the influence of the English language on German communicative preferences. Towards a model for the analysis of language use in visual media." (Doctoral dissertation, University of Hamburg, Germany.)

Bernolet, S., Hartsuiker, R. J. and Pickering, M. J. 2007. "Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition." *Journal of Experimental Psychology. Learning, Memory, and Cognition* 33 (5): 931.

Caramazza, A. and Brones, I. 1979. "Lexical access in bilinguals." *Bulletin of the Psychonomic Society* 13: 212–214.

Catford, J. C. 1965. *A Linguistic Theory of Translation*. Oxford: Oxford University Press.

Chesterman, A. 1997. *Memes of Translation*. Amsterdam/Philadelphia: John Benjamins.

Christoffels, I. K., de Groot, A. M. B. and Kroll, J. F. 2006. "Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency." *Journal of Memory and Language* 54 (3): 324–345.

Dijkstra, T., Timmermans, M. and Schriefers, H. 2000. "On being blinded by your other language: Effects of task demands on interlingual homograph recognition." *Journal of Memory and Language* 42 (4): 445–464.

Dowty, D. 1991. "Thematic proto-roles and argument selection." *Language* 67: 547–619.

Duyck, W., Desmet, T., Verbeke, L. P. C. and Brysbaert, M. 2004. "WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French." *Behavior Research Methods, Instruments, & Computers* 36 (3): 488–499.

Duyck, W., Van Assche, E., Drieghe, D. and Hartsuiker, R. J. 2007. "Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access." *Journal of Experimental Psychology. Learning, Memory, and Cognition* 33 (4): 663.

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process*. Amsterdam/Philadelphia: John Benjamins.

Gambier, Y. 2008. "Stratégies et tactiques en traduction et interprétation." In *Efforts and Models in Interpreting and Translation Research. A tribute to Daniel Gile*, G. Hansen, A. Chesterman and H. Gerzymisch-Arbogast (eds), 63–82. Amsterdam/Philadelphia: John Benjamins.

Göpferich, S. 2000. "Textkritik als Mittel zur Verbesserung der Textproduktionskompetenz angehender Technischer Redakteure, vorgeführt am Beispiel einer Kundeninformationsbroschüre der Volkswagen AG." *Fachsprache/International Journal of LSP* 3 (4): 106–132.

———. 2004. "Sprachwissenschaftliche Übersetzungsanalysen und Übersetzungsvergleiche am Beispiel von technischen Texte." In *Übersetzung. Translation. Traduction. Ein internationales Handbuch zur Übersetzungsforschung. An International Encyclopedia of Translation Studies. Encyclopédie internationale de la recherche sur la traduction.* [*Handbücher zur Sprach- und Kommunikationswissenschaft* 26.1], H. Kittel, A. P. Frank, N. Greiner, T. Hermans, W. Koller, J. Lambert and F. Paul (eds), 719–727. Berlin/New York: Walter de Gruyter.

———. 2008. *Translationsprozessforschung. Stand – Methoden – Perspektiven* [*Translationswissenschaft* 4]. Tübingen: Gunter Narr.

———. 2009. "Towards a model of translation competence and its acquisition: The longitudinal study TransComp." In *Behind the Mind. Methods, Models and Results in Translation Process Research* [CBS in Language 37], S. Göpferich, A. L. Jakobsen and I. Mees (eds), 11–38. Copenhagen: Samfundslitteratur.

Hartsuiker, R. J., Pickering, M. J. and Veltkamp, E. 2004. "Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals." *Psychological Science* 15 (6): 409–414.

Hawkins, J. A. 1986. *A comparative typology of English and German*. London/Sydney: Croom Helm.

Keuleers, E., Brysbaert, M. and New, B. (under revision). "Subtlex-nl: A new frequency measure for Dutch words based on film subtitles." *Behavior Research Methods*.

Kucera, H. and Francis, W. N. 1967. *Computational Analysis of Present-Day American English*. Providence, R.I.: Brown University Press.

Lee, Y.-O. 2001. "Musayngmul cwue thatongsakwumunuy yenghanpenyek. [English-Korean translation of non-human subject-transitive verb constructions]." *Journal of Translation Studies* 2 (1): 53–76.

Libben, M. R. and Titone, D. A. 2009. "Bilingual lexical access in context: Evidence from eye movements during reading." *Journal of Experimental Psychology: Learning, Memory and Cognition* 35 (2): 381–391.

Master, P. 1991. "Active verbs with inanimate subjects in scientific prose." *English for Specific Purposes* 10: 15–33.

Mauranen, A. and Kujamäki, P. (eds). 2004. *Translation Universals: Do They Exist?* Amsterdam/Philadelphia: John Benjamins.

McClelland, J. L. and Rumelhart, D. E. 1981. "An interactive activation model of the effect of context in perception. Part I: An account of basic findings." *Psychological Review* 88: 375–407.

Nida, E. A. 1964. *Toward a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating*. Leiden: E. J. Brill.

Nord, C. 1991. *Text Analysis in Translation. Theory, Method, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Amsterdam/Atlanta, Ga.: Rodopi.

PACTE Group. 2000. "Acquiring translation competence: Hypotheses and methodological problems of a research project." In *Investigating Translation: Selected Papers from the 4th International Congress on Translation, Barcelona 1998*, A. Beeby, D. Ensinger and M. Presas (eds), 99–106. Amsterdam/Philadelphia: John Benjamins.

———. 2009. "Applying the scientific method to translation studies: problems and solutions in research into translation competence." Paper presented at the International Research Workshop Methodology in Translation Process Research, University of Graz, April 2009, 6–8.

Ruiz, C., Paredes, N., Macizo, P. and Bajo, M. T. 2008. "Activation of lexical and syntactic target language properties in translation." *Acta Psychologica* 128 (3): 490–500.

Rumelhart, D. E. and McClelland, J. L. 1985. "Levels indeed! A response to Broadbent." *Journal of Experimental Psychology: General* 114 (2): 193–197.

Schoonbaert, S., Hartsuiker, R. J. and Pickering, M. J. 2007. "The representation of lexical and syntactic information in bilinguals. Evidence from syntactic priming." *Journal of Memory and Language* 56: 153–171.

Schoonbaert, S., Duyck, W., Brysbaert, M. and Hartsuiker, R. J. 2009. "Semantic and translation priming from a first language to a second and back: Making sense of the findings." *Memory & Cognition* 37 (5): 569–586.

Shlesinger, M. 1998. "Corpus-based interpreting studies as an offshoot of corpus-based translation studies." *Meta* XLIII (4): 468–493.

Slabakova, R. and Montrul, S. 2002. "On aspectual shifts in L2 Spanish." In *Proceedings of the Boston University Conference on Language Developmen*t, B. Skarabela, S. Fish and A.-H. Do (eds), 631–642. Somerville, Mass.: Cascadilla Press.

Sunderman, G. and Schwartz, A. I. 2008. "Using cognates to investigate cross-language competition in second language processing." *TESOL Quarterly* 42: 527–536.

Thome, G. 2004. "Typologie der Übersetzungsschwierigkeiten aus sprachwissenschaftlicher Sicht." In *Übersetzung. Translation. Traduction. Ein internationales Handbuch zur Übersetzungsforschung. An International Encyclopedia of Translation Studies. Encyclopédie internationale de la recherche sur la traduction* [*Handbücher zur Sprach- und Kommunikationswissenschaft* 26.1], H. F. Kittel, Armin Paul/Greiner, Norbert/Hermans, Theo/Koller, Werner/Lambert, José/Paul, Fritz (eds), 436–446. Berlin/New York: de Gruyter.

Toury, G. 1995. *Descriptive Translation Studies – and Beyond* [Benjamins Translation Library 4]. Amsterdam/Philadelphia: John Benjamins.

Van Assche, E., Duyck, W., Hartsuiker, R. J. and Diependaele, K. 2009. "Does bilingualism change native-language reading? Cognate effects in a sentence context." *Psychological Science* 20 (8): 923–927.

Vandepitte, S. 2007. "Semantic and pragmatic meanings in translation." *Belgian Journal of Linguistics. The Study of Language and Translation* 21: 185–200.
van Hell, J. G. 2002. "Bilingual word recognition beyond orthography: On meaning, linguistic context and individual differences." *Bilingualism* 5 (3): 175–224.
van Hell, J. G. and de Groot, A. M. B. 2008. "Sentence context modulates visual word recognition and translation in bilinguals." *Acta Psychologica* 128 (3): 431–451.
Weber, A. and Cutler, A. 2004. "Lexical competition in non-native spoken-word recognition." *Journal of Memory and Language* 50 (1): 1–25.
Whaley, C. P. 1978. "Word-nonword classification time." *Journal of Verbal Learning and Verbal Behavior* 17: 143–154.

Appendix A. Questionnaire (translation from Dutch)

Answer these questions as truthfully as possible.
There are no correct or wrong answers: the answers only reflect your situation.

What is your mother tongue?
How many years have you studied English (minimum two hours per week)?
How old were you when you started to learn English?
How many months have you spent in an English speaking country?
Which other languages do you know?
Underline the language that you speak at home:
dialect, standardized dialect ('tussentaal'), standard
If any other languages, which?
How do you assess your own competences in English? Tick the relevant box:

	Moderate	Good	Very good
Listening	O	O	O
Reading	O	O	O
Speaking	O	O	O
Writing	O	O	O
Translating into Dutch	O	O	O

How many years have you typed on a computer?
Do you touch type?
What is your typing speed?
Is the keyboard in front of you the same that you usually use?
Your sight is: O normal O corrected to normal (with glasses or lenses)
O weakened

Is there anything you want to say about this experiment?

Appendix B. Stimuli

	Critical words		Control words	
	EN	NL targets	EN	NL targets
1	action	actie	actor	acteur
2	address	adres	actress	actrice
3	advice	advies	angel	engel
4	barrier	barrière	athlete	atleet
5	career	carrière	author	auteur
6	climate	klimaat	banker	bankier
7	culture	cultuur	bishop	bisschop
8	debate	debat	brother	broer
9	defence	defensie	captain	kapitein
10	essence	essentie	champion	kampioen
11	function	functie	clerk	klerk
12	harmony	harmonie	client	klant
13	ideal	ideaal	comrade	kameraad
14	justice	justitie	couple	koppel
15	manner	manier	cousin	kozijn
16	minute	minuut	criminal	crimineel
17	mission	missie	dancer	danser
18	music	muziek	daughter	dochter
19	mystery	mysterie	doctor	dokter
20	nature	natuur	father	vader
21	number	nummer	friend	vriend
22	passion	passie	governor	gouverneur
23	pattern	patroon	guest	gast
24	pension	pensioen	leader	leider
25	period	periode	master	meester
26	plague	plaag	mother	moeder
27	practice	praktijk	musician	muzikant
28	process	proces	officer	officier
29	profit	profijt	patient	patiënt
30	revolt	revolte	person	persoon
31	season	seizoen	priest	priester
32	section	sectie	public	publiek
33	standard	standaard	sister	zuster
34	success	succes	slave	slaaf
35	summer	zomer	soldier	soldaat
36	symptom	symptoom	speaker	spreker
37	theory	theorie	thief	dief
38	triumph	triomf	tourist	toerist
39	version	versie	winner	winnaar
40	vision	visie	worker	werker

Appendix C. Pre-test word translation onset time

Non-proto-typical agent / proto-typical agent		Untrained students			Trained students			All students		
		Total onset time / word	Mean onset time / word	No	Total onset time / word	Mean onset time / word	No	Total onset time / word	Mean onset time / word	No
N	ation	27.22	1.05	26	24.99	0.93	27	52.21	0.99	53
N	address	31.69	1.13	28	31.82	1.18	27	63.51	1.15	55
N	advice	33.81	1.21	28	28.37	1.05	27	62.18	1.13	55
N	barrier	30.32	1.44	21	29.43	1.34	22	59.75	1.39	43
N	career	33.7	1.25	27	27.89	1.07	26	61.59	1.16	53
N	climate	33.49	1.2	28	28.17	1.04	27	61.66	1.12	55
N	culture	30.34	1.12	27	27.48	1.02	27	57.82	1.07	54
N	debate	33.33	1.23	27	25.77	0.95	27	59.1	1.09	54
N	defence	41.37	1.65	25	34.5	1.38	25	75.87	1.52	50
N	essence	35.19	1.41	25	31.49	1.31	24	66.68	1.36	49
N	function	27.37	1.05	26	25.64	0.95	27	53.01	1	53
N	harmony	33.99	1.26	27	31.43	1.16	27	65.42	1.21	54
N	ideal	29.99	1.25	24	31.81	1.33	24	61.8	1.29	48
N	justice	37.65	1.71	22	42.07	1.75	24	79.72	1.73	46
N	manner	31.5	1.31	24	34.09	1.31	26	65.59	1.31	50
N	minute	30.38	1.22	25	25.52	1.02	25	55.9	1.12	50
N	mission	28.63	1.1	26	27.96	1.04	27	56.59	1.07	53
N	music	28.24	1.05	27	26.32	1.01	26	54.56	1.03	53
N	mystery	28.89	1.11	26	29.83	1.15	26	58.72	1.13	52
N	nature	30.12	1.16	26	26.42	1.02	26	56.54	1.09	52
N	number	28.33	1.18	24	29.22	1.08	27	57.55	1.13	51
N	passion	29.17	1.08	27	25.76	0.95	27	54.93	1.02	54
N	pattern	35.8	1.38	26	29.68	1.1	27	65.48	1.24	53
N	pension	25.02	1.32	19	29.46	1.34	22	54.48	1.33	41
N	period	33.46	1.24	27	26.29	1.05	25	59.75	1.15	52
N	plague	31.14	1.42	22	30.25	1.12	27	61.39	1.25	49
N	practice	40.48	1.69	24	29.14	1.21	24	69.62	1.45	48
N	process	30.25	1.16	26	24.58	0.98	25	54.83	1.08	51
N	profit	40.68	1.56	26	34.81	1.34	26	75.49	1.45	52
N	revolt	29.15	1.82	16	39.1	1.86	21	68.25	1.84	37
N	season	32.6	1.16	28	25.82	1.03	25	58.42	1.1	53
N	section	32.51	1.2	27	27.63	1.11	25	60.14	1.16	52
N	standard	30	1.15	26	26.84	1.03	26	56.84	1.09	52
N	success	27.59	0.99	28	22.36	0.86	26	49.95	0.93	54
N	summer	32.39	1.2	27	25.56	1.02	25	57.95	1.11	52
N	symptom	33.57	1.2	28	21.15	0.88	24	54.72	1.05	52
N	theory	31.03	1.11	28	23.2	0.86	27	54.23	0.99	55
N	triumph	33.14	1.33	25	31.4	1.16	27	64.54	1.24	52
N	version	32.1	1.19	27	22.6	0.9	25	54.7	1.05	52
N	vision	33.05	1.27	26	32.09	1.34	24	65.14	1.3	50

Non-proto-typical agent / prototypical agent		Untrained students			Trained students			All students		
		Total onset time / word	Mean onset time / word	No	Total onset time / word	Mean onset time / word	No	Total onset time / word	Mean onset time / word	No
P	actor	28.46	1.02	28	23.79	0.88	27	52.25	0.95	55
P	actress	35.08	1.25	28	26.43	1.02	26	61.51	1.14	54
P	angel	32.6	1.21	27	29.15	1.12	26	61.75	1.17	53
P	athlete	33.48	1.2	28	27.96	1.04	27	61.44	1.12	55
P	author	33.2	1.23	27	26.27	1.01	26	59.47	1.12	53
P	banker	42.3	1.63	26	35.21	1.3	27	77.51	1.46	53
P	bishop	33.57	1.34	25	26.5	1.02	26	60.07	1.18	51
P	brother	30.13	1.12	27	27.11	1	27	57.24	1.06	54
P	captain	34.44	1.23	28	26.42	1.02	26	60.86	1.13	54
P	champion	33.31	1.28	26	32.58	1.21	27	65.89	1.24	53
P	clerk	31.81	1.99	16	33.43	1.59	21	65.24	1.76	37
P	client	35.39	1.36	26	34.01	1.31	26	69.4	1.33	52
P	comrade	30.47	1.45	21	35.5	1.48	24	65.97	1.47	45
P	couple	37.67	1.35	28	29.41	1.23	24	67.08	1.29	52
P	cousin	38.81	1.62	24	35.08	1.46	24	73.89	1.54	48
P	criminal	36.61	1.41	26	27.72	1.11	25	64.33	1.26	51
P	dancer	34.21	1.22	28	27.55	1.06	26	61.76	1.14	54
P	daughter	28.37	1.05	27	24.56	0.91	27	52.93	0.98	54
P	doctor	29.89	1.11	27	28.06	1.04	27	57.95	1.07	54
P	father	33.71	1.2	28	26.39	1.02	26	60.1	1.11	54
P	friend	28.3	1.01	28	26.02	1	26	54.32	1.01	54
P	governor	38.54	1.48	26	28.03	1.12	25	66.57	1.31	51
P	guest	29.35	1.13	26	27.37	1.01	27	56.72	1.07	53
P	leader	31.29	1.16	27	26.6	0.99	27	57.89	1.07	54
P	master	40.68	1.63	25	26.17	1.25	21	66.85	1.45	46
P	mother	32.98	1.18	28	22.27	0.86	26	55.25	1.02	54
P	musician	36.56	1.31	28	27.22	1.09	25	63.78	1.2	53
P	officer	34.12	1.36	25	28.28	1.35	21	62.4	1.36	46
P	patient	36.6	1.31	28	28.02	1.04	27	64.62	1.17	55
P	person	28.91	1.07	27	29.9	1.15	26	58.81	1.11	53
P	priest	37.19	1.43	26	29.25	1.13	26	66.44	1.28	52
P	public	32.87	1.22	27	32.12	1.24	26	64.99	1.23	53
P	sister	32.08	1.23	26	29.3	1.13	26	61.38	1.18	52
P	slave	27.92	1	28	25.13	0.97	26	53.05	0.98	54
P	soldier	29.24	1.12	26	26.85	1.03	26	56.09	1.08	52
P	speaker	36.73	1.47	25	30.54	1.27	24	67.27	1.37	49
P	thief	30.08	1.11	27	22.43	0.93	24	52.51	1.03	51
P	tourist	26.76	0.99	27	25.78	0.95	27	52.54	0.97	54
P	winner	28.83	1.11	26	27.71	1.03	27	56.54	1.07	53
P	worker	39.59	1.47	27	34.68	1.39	25	74.27	1.43	52
	Grand Total	2551.9	1.27	28	2227.94	1.13	27	4779.84	1.2	55

Appendix D1. English critical words versus control words (pre-test)

Critical words	Length	Frequency	Summated bigram	Control words	Length	Frequency	Summated bigram
action	6	2.3802	16298	actor	5	1.9345	8321
summer	6	2.1072	11382	actress	7	1.2788	16632
pension	7	1.7324	17627	angel	5	1.3802	11298
mission	7	1.5563	18139	athlete	7	1.2553	17295
career	6	1.8573	19726	author	6	1.6128	8846
profit	6	1.8865	8936	banker	6	1.1761	14096
ideal	5	1.7782	10363	bishop	6	1.5185	8705
period	6	2.3201	15998	brother	7	2.1614	15744
season	6	1.8451	13247	captain	7	1.8692	14436
vision	6	1.7404	14526	champion	8	1.2788	14809
defence	7	2.0934	12087	clerk	6	1.5185	12517
manner	6	1.9031	17661	client	6	1.415	12586
nature	6	2.2856	13586	comrade	7	1.2304	12053
function	8	1.9731	18671	couple	6	2.1335	12102
barrier	7	1.4771	17199	cousin	6	1.5315	17387
version	7	1.6812	20549	criminal	8	1.6232	19551
music	5	2.1139	7717	dancer	6	1.1761	15884
culture	7	1.9191	9781	daughter	8	2.1271	14888
symptom	7	1.3979	5141	doctor	6	2.2765	8212
advice	6	1.8573	7325	father	6	2.5786	16324
justice	7	1.7482	16847	friend	6	2.5635	11467
theory	6	2.1553	8372	governor	8	1.415	15804
success	7	2.0607	10542	guest	6	1.7993	8971
triumph	7	1.5441	7766	leader	6	2.1584	17925
minute	6	2.4829	15074	master	6	1.8808	19932
pattern	7	2.0969	19436	mother	6	2.7168	13345
plague	6	1.1139	5239	musician	8	1.1761	13594
debate	6	1.7559	13483	officer	7	2.1614	14874
number	6	2.6021	9984	patient	7	1.8808	20129
process	7	2.3054	14150	person	6	2.4065	16274
passion	7	1.6128	16368	priest	6	1.7076	13862
address	7	1.9345	13738	public	6	2.5527	8976
climate	7	1.5315	16186	sister	6	2.0719	21512
practice	8	2.2122	18403	slave	5	1.5682	6153
section	7	2.0792	18608	soldier	7	1.9494	10000
revolt	6	1.1761	7657	speaker	7	1.415	14050
mystery	7	1.5051	17319	thief	5	1.0792	4648
essence	7	1.2788	17705	tourist	7	1.5682	17631
standard	8	2.1173	18670	winner	6	1.1761	19074
harmony	7	1.1761	12891	worker	6	2.3284	12546
Sum	**263**	**74.3944**			**256**	**70.6513**	
Variance	**1604.2**	**0.1332**	**19081709**		**0.708**	**0.2164029**	**16126820.5**

t-value length 0.3185 ***t-value freq*** 0.3202 ***t-value bigram*** 0.9588

Appendix D2. Dutch critical words versus control words with near cognate translations (pre-test)

Critical words	Length	Frequency	Summated bigram	Control words	Length	Frequency	Summated bigram
actie	5	1.8062	35569	gast	4	1.7482	25703
zomer	5	1.9085	54129	dief	4	1.1139	25502
visie	5	1.7243	31379	broer	5	2.1072	61156
plaag	5	0.4771	30906	engel	5	1.415	82881
triomf	6	1.0792	24295	klant	5	1.699	34332
debat	5	1.1761	27793	vader	5	2.7604	59919
adres	5	1.6990	27877	slaaf	5	1.6435	30002
missie	6	1.0792	36336	vriend	6	2.4533	76260
ideaal	6	1.8195	43619	klerk	5	0.4771	56926
muziek	6	2.0607	27934	acteur	6	1.2553	33821
natuur	6	1.9685	21684	atleet	6	0.6021	37303
manier	6	2.5740	82994	auteur	6	1.7709	31088
minuut	6	2.2405	27937	danser	6	0.6021	68933
nummer	6	1.8573	52055	dokter	6	2.1614	65421
proces	6	2.2122	23812	koppel	6	0.699	36285
passie	6	0.8451	32333	kozijn	6	0.4771	19001
succes	6	1.9685	11822	leider	6	1.8513	79738
sectie	6	1.1461	40019	moeder	6	2.7752	71463
advies	6	1.6990	31120	werker	6	1.6021	98137
versie	6	1.3222	77207	zuster	6	2.017	81462
profijt	7	0.6021	28037	actrice	7	0.8451	29494
periode	7	2.1271	76695	bankier	7	1.0414	81173
cultuur	7	1.9956	13379	dochter	7	2.0792	84476
seizoen	7	1.3010	59549	meester	7	1.8865	104107
functie	7	2.1987	34980	patiënt	7	2.1072	84237
patroon	7	1.6721	49266	persoon	7	2.29	76147
revolte	7	0.3010	47040	publiek	7	2.017	38702
theorie	7	2.0864	49813	soldaat	7	2.0531	35979
klimaat	7	1.4624	44638	spreker	7	1.1761	78025
carrière	8	1.3222	97976	toerist	7	1.3424	95254
symptoom	8	1.4150	23608	winnaar	7	0.6021	54322
pensioen	8	1.1139	94702	bisschop	8	1.5315	47644
barrière	8	0.9542	97976	kameraad	8	1.5682	84631
mysterie	8	1.1761	108776	kampioen	8	0.6021	56846
essentie	8	1.0000	88336	kapitein	8	1.9956	63348
defensie	8	1.0414	78351	officier	8	1.8451	66905
justitie	8	1.3617	68900	priester	8	1.6628	119673
harmonie	8	1.3423	55893	muzikant	8	0.7782	32365
praktijk	8	2.0086	44806	crimineel	9	0.699	74984
standaard	9	0.9542	83004	gouverneur	10	1.2788	72537
Sum	**266.0000**	**60.0992**	**1986545**		**262**	**60.6325**	**2456182**
Variance	**1.2077**	**0.2849**	**685854666**		**1.6897**	**0.4133204**	**625807469**

t-test length 0.711259 ***t-test freq*** 0.919886 ***t-test bigram*** 0.043693

Appendix E. Main test materials

This success raised expectations for the new performance
This actor raised expectations for the new performance

His theory contradicted what I suspected
His daughter contradicted what I suspected

Most of the action came before Friday afternoon
Most of the tourists came before Friday afternoon

That function worked for me
That slave worked for me

His passion found a home in our house
His friend found a home in our house

Music took him around the world
Mother took him around the world

The symptom took care of itself
The thief took care of himself

My version of the test obtained the teacher's reactions to the pupils
My brother obtained the teacher's reactions to the pupils

This culture changed the city
This doctor changed the city

The mission developed a new strategy
The guest developed a new strategy

The process determined the number of people
The leader determined the number of people

Such a nature rendered the neighbourhood unsafe
Such a soldier rendered the neighbourhood unsafe

This debate affected the country
This winner affected the country

Such a standard saved a lot of money
Such a father saved a lot of money

The first season put us in the most difficult group
The first person put us in the most difficult group

That summer took us to Texas
That athlete took us to Texas

This climate placed many people under stress
This author placed many people under stress

That minute inspired many people
That captain inspired many people

Their advice raised a difficult matter
Their actress raised a difficult matter

Mysteries surrounded him
Dancers surrounded him

The licence number told me he was from France originally
The angel told me he was from France originally

That period put much emphasis on metaphors
That bishop put much emphasis on metaphors

His address changed nothing
His patient changed nothing

That section described the different parts of the violin
That musician described the different parts of the violin

His career focussed on economics
His sister focussed on economics

This harmony brought them peace
This public brought them peace

This pattern has now returned for the first time
This champion has now returned for the first time

This triumph brought the two together
This priest brought the two together

This ideal carried the company forward
This couple carried the company forward

This vision wove together political and economic matters
This governor wove together political and economic matters

These manners earned her respect
These clients earned her respect

This pension only bought cheap things
This officer only bought cheap things

These barriers protected us against enemies
These workers protected us against enemies

The best practice combined two approaches
The best master combined two approaches

This profit undermined customers' trust
This banker undermined customers' trust

His defence put me off a little
His comrade put me off a little

Plague visited the area
The criminal visited the area

Man's essence tells us how little we know
The speaker tells us how little we know

Sight translation and speech disfluency

Performance analysis as a window to cognitive translation processes

Gregory M. Shreve, Isabel Lacruz and Erik Angelone
Kent State University, USA

Written and sight translation share a comprehension component centered on written input. In sight translation, because of production constraints, the cognitive effort expended in a given span of time is greater than in written translation. Comprehension, transfer, and production processes occur in much shorter periods of time, and this *compression* of the information processing window has marked consequences for the target production. Translation problems encountered during sight translation are likely to be more *disruptive* to translation processing. Disruptions will manifest not just as errors or deficiencies in rendering, but as *speech disfluencies* in the oral performance of the translation. In this paper, we propose that the analysis of speech disfluencies occurring during sight translation performance provides significant information about cognitive phenomena associated with sight translation such as *visual interference*, as well as about cognitive processes associated with the solution of lexical, syntactic, and strategic translation problems.

Keywords: sight translation, speech disfluency, cognitive effort, visual interference, translation problem-solving

1. Introduction

Written translation and sight translation share a comprehension component centered on written input. However, in sight translation, because of real-time performance and concurrent production constraints, the comprehension and transfer effort expended in a given span of time is likely to be greater than in written translation. Reading, comprehension, transfer, and production must occur in much shorter periods of time, and this *compression* of the information processing window will have visible, marked consequences on the output. Any translation problems encountered during the sight translation process are likely to be more *disruptive* to translation

processing and consequently have a greater impact on the translation quality of the output than they would in written translation. Because output in sight translation must be *verbalized*, disruptions will manifest not just as errors or deficiencies in rendering, but also as speech disfluencies in the oral performance of the translation. We propose that analysis and classification of the target text speech disfluencies resulting from the time and verbalization constraints of the sight translation task can provide significant information about underlying cognitive processes.

If compared to written translation, the constraints of sight translation allow much less time for optimal execution of the processes associated with rendering a target text. In written translation, the translator can potentially allocate cognitive resources to comprehension, transfer, and production processes as needed, thereby *distributing* the translation effort over longer spans of time more strategically and *de-emphasizing* the visible effects in the target text of difficulties encountered in the process of rendering the source text. More time can potentially be allowed for constituent processes of translation, such as consideration of translational cues in the source text, translation problem recognition and evaluation, solution proposal, thinking through alternative problem-solving strategies, or even core operations such as accessing lexical memory and performing syntactic shifts. The time compression of the sight translation task *emphasizes* and therefore reveals, because it forces disruptions and errors, aspects of underlying cognitive states and phenomena that are less visible or marked in the written translation products.

There are at least two product-oriented consequences of the compression of the temporal window in sight translation. First, we can expect increases in the number and shifts in the nature of errors detectable in the sight translation as compared to the written translation. Analysis of differentials in the frequency, distribution, location, and type of errors between sight and written translation can help highlight details of some of the underlying cognitive processes more clearly because the effects of the repair and revision allowed during written production are removed. We do not discuss this first consequence in this paper.

A second, and perhaps very interesting consequence of time compression, is a class of events we discovered when we applied a detailed analytical protocol to screen recordings of sight translations performed by participants in a sight translation experiment. We discovered that error patterns, while important, were not the only interesting or revealing phenomena contained in transcripts of the screen recordings. An entire array of *speech disfluencies* (hereafter DF) emerged in the recorded performance data. Following Gósy (2007: 93), we define speech disfluencies as unintentional "phenomena that interrupt the flow of speech and do not add propositional content to an utterance."

Even if disfluencies are unintentional, this does not mean that they are random speech occurrences. A number of studies in the psycholinguistic disfluency literature have argued that disfluencies are associated with underlying cognitive states or processes. The main thrust of the literature argues that an important reason that speech disfluencies occur is because speakers are having language production problems

(Clark and Wasow 1998). Corley and Stewart (2008) summarize the literature on the cognitive implications of disfluencies in language production, citing studies that associate them with situations of increased cognitive effort (Barr 2001; Bortfeld, Bloom, Schober and Brennan 2001; Merlo and Mansur 2004), and with cases where there is uncertainty about what to say next, as for instance when there is an element of choice in selecting words (Beattie and Butterworth 1979; Schachter, Christenfeld, Ravina and Bilous 1991). The cognitive effort can derive from a number of underlying situations. For instance, Fox Tree and Clark (1997) and Schnadt and Corley (2006) have associated speech disfluencies with planning problems in production (semantic planning, syntactic planning, and lexical planning). Other studies have noted the association of disfluencies with lexical retrieval, lexical access, and, particularly, uncertainty in lexical choice (Levelt 1983; Schachter, Christenfeld, Ravina and Bilous 1991).

Because speech disfluencies have been shown in the psycholinguistic literature to be predictors of cognitive effort due to lexical or syntactic uncertainty, planning, and production problems, we felt that analysis of these speech phenomena might offer a unique opportunity for translation researchers to understand the spoken translation product as a direct reflection of underlying cognitive processing. They should be relatively strong indicators of processing effort associated with translation problem recognition and resolution. While the study of speech disfluencies in language mediation is not a completely new research area, since scholars of simultaneous interpreting have also studied such disfluencies (cf. Pöchhacker 1995), we believe that ours is the first report on speech disfluencies in sight translation. Our main purpose in this paper is to show that disfluency performance analysis, especially if triangulated with another form of data such as eye tracking, can contribute to our interpretation of how the recognition and resolution of sight translation problems plays out cognitively.

2. Method: Analysis of speech disfluencies in the sight translation protocols

Our research design used a syntactic manipulation involving a set of two-paragraph texts where one text version (the *A-Text*) contained a syntactically complex area of interest (AOI) in the second paragraph and a non-complex AOI in the first. In the other version (the *B-Text*), a complex AOI was placed first and a non-complex AOI second. We chose a syntactic manipulation which has been shown to produce eye tracking effects, and we assumed it might also produce other indicators of cognitive effort such as errors and speech disfluencies. Experimental participants translated two texts, one *A-Text* and one *B-Text*, using sight translation. They then translated an *A-Text* and a *B-Text* using written translation. Each participant performed both tasks (sight and written), but performed only one such task on any given text. In the design, the independent variable was syntactic complexity. The dependent variables were eye tracking metrics and errors and disfluencies in verbal output for the sight translation, and keystroke log patterns, pauses, and errors in written output for the written translation.

Our initial hypotheses predicted that:

- For sight translation, as the syntactic complexity of a segment of a source text increases, comprehension, transfer, and production become correspondingly more difficult and disrupted. That disruption may be measured by an increased number of output errors (including disfluency phenomena) and by eye tracking metrics indicating greater processing effort, e.g., greater numbers of fixations, longer fixation durations, and greater numbers of regressive fixations (regressions).
- For written translation, as the syntactic density of the source text segment increases, both comprehension and production effort, measured by data from keystroke logs (e.g., pauses, pause durations, cursor movements) and errors in the target text, will be less significantly disrupted due to the ability to reallocate effort over longer periods of time and the reduced sensitivity to source text disruption.

Details of the eye tracking results of the experiment as well as a comparison of sight translation and written translation results are reported and discussed in detail in Shreve, Lacruz and Angelone (2010). In the present paper we will report and discuss sight translation disfluency results relative to our first hypothesis.

2.1 Experimental procedure

Participants (see Section 2.3) were tested individually. Each participant was seated in front of a computer, given the same set of instructions, and then calibrated to the eye tracking device, a Tobii© 1750 eye tracker. Each was then instructed to read a general language "practice text" in Spanish. The practice Spanish text was similar in topic, length, and difficulty to the four experimental texts. After the practice text was read, each participant then translated two general language Spanish texts as sight translation tasks. Participants were first given an *A-text* to translate, and when they were done with it, they were given a chance to rest for a few minutes before the second text, a *B-text*, appeared on the screen. As indicated, all participants translated two different texts, one in version A and one in version B. As a result, each participant sight translated two complex paragraphs and two non-complex paragraphs. Sessions were not timed, and each session lasted approximately thirty minutes.

2.2 Experimental texts

For the sight translation and written translation experiments, four general language Spanish texts were used. All texts were adapted from past accreditation exams from the American Translators Association, and all were about the topic of tourism. The four texts used in the sight translation and written translation had a mean length of 167 words, ranging from 159 to 171 words. Each text consisted of two paragraphs of similar length.

A–TEXT

NC

El 12 de junio de 1926 parecía que la mitad de Barcelona estaba guardando 1uto. Había un cortejo fúnebre de aproximadamente un kilómetro. Este cortejo se dirigía con lentitud desde el hospital de Santa Cruz, en la ciudad antigua, hacia la famosa Sagrada Familia. Millares de personas se detenían en las calles para rendirle el último honor a Antoni Gaudí. El pintor uruguayo Joaquín Torres García le llamó « el arquitecto más genial, el más catalán entre los catalanes ». Casi todos los altos dignatarios de la región tomaron parte en el cortejo fúnebre.

C

Como Gaudí se había convertido desde hacia tiempo en uno de los héroes populares, el gobierno que había obtenido permiso del Papa que no pudo asistir personalmente, ordenó que su cadáver fuera depositado en la iglesia inconclusa. Gaudí encontró el último reposo en el lugar donde había trabajado 43 años de su vida y al que había dedicado sus 12 últimos años en exclusiva. Había creado su propio espacio personal donde se le dedicó un glorioso sepulcro.

B–TEXT

C

El 12 de junio de 1926 parecía que la mitad de Barcelona estaba guardando luto. Había un cortejo fúnebre de aproximadamente un kilómetro. Este cortejo, el cual estaba rodeado de millares de personas detenidas en las calles para rendirle el último honor a Antoni Gaudí, se dirigía desde el hospital de Santa Cruz, en la ciudad antigua, hacia la Sagrada Familia. El pintor uruguayo Joaquín Torres García le llamo « el arquitecto mas genial, el mas catalán entre los catalanes ». Casi todos los altos dignatarios de la región tomaron parte en el cortejo fúnebre.

NC

Gaudí se había convertido desde hacia tiempo en un héroe popular. El gobierno había pedido permiso para que su cadáver fuera depositado en la iglesia inconclusa. El Papa dio su conformidad pero no pudo asistir al entierro. Gaudí encontró el ultimo reposo en el lugar donde había trabajado 43 años de su vida y al que había dedicado sus 12 últimos años en exclusiva. Había creado su propio espacio personal donde se le dedicó un glorioso sepulcro.

Figure 1. Experimental stimulus texts

As indicated, one paragraph in each text was manipulated to create two text versions, designated A and B. The length, vocabulary, and meaning were identical for both versions of each paragraph, except for a syntactic manipulation of a single segment of the paragraph that was designated as an AOI (Area of Interest). The AOIs were all the same length, approximately thirty-six words. The syntactically non-complex AOIs were made up of simple, short sentences. The syntactically complex version was modified to contain several embedded relative clauses. Operational definitions of *syntactically complex* and *syntactically non-complex* were informal. Figure 1 shows a sample set of experimental stimulus texts with the complex and non-complex AOIs indicated in the outlined areas.

2.3 Participants

Eleven participants took part in the study. The participant pool was derived from the student body of the graduate translation programs at Kent State University. Seven were in the second semester and four in the fourth semester in the M.A. in Translation

program, enrolled in the Spanish-English track. One was a doctoral student. Five of the participants had some previous professional experience, but no participant had experience exceeding two years. The mean age was twenty-six, and there were seven females and four males. Ten of the participants were native speakers of English, and one was a native speaker of Spanish. The native speakers of English had studied Spanish in a formal educational setting for an average of eight years. The native speaker of Spanish studied English in a formal educational setting for ten years.

2.4 Eye tracking measures

Eye tracking metrics such as numbers of fixations, longer fixation durations, and greater numbers of regressive fixations (regressions) have been shown to be sensitive to syntactic, semantic, and pragmatic factors that affect sentence processing (Rayner, Sereno, Morris, Schmauder and Clifton 1989; Rayner 1998; Traxler, Morris and Seely 2002). Kemper, Crow and Kemtes (2004) and others have also established that increases in regressions and in regression times are associated with difficulty in parsing sentences. Thus, we had every expectation on theoretical grounds to expect a syntactic effect that would be expressed in eye movement data indicative of increased cognitive effort.

2.5 Disfluency measures

For the purposes of this paper, we have adopted Gósy's definition as a broad starting point for classifying disfluent speech phenomena in the sight translations. There are a number of different approaches to both the terminology and classification of disfluencies in the psycholinguistic literature (for a discussion, see Shriberg 1994). From these approaches, we have developed a scheme that appears to represent the range of these phenomena in the corpus of twenty transcripts we studied. The four categories of disfluencies discussed below and the terminology we use are well documented in the literature. We have not included definitions of all possible disfluencies, but only those that occurred with any frequency in our corpus. The examples given are from the sight translations.

2.6 Unfilled pause

Unfilled or silent pauses are interruptions in speech that do not contain any phonetic content. For further discussion see Goldman-Eisler (1968) and Shriberg (1994). Unfilled pauses are assumed to be employed either intentionally or automatically for the purpose of delaying speech production. For the purposes of this analysis, we defined an unfilled pause as any pause with duration greater than or equal to one second. This floor was selected so as to allow the later comparison of the sight translation pause data with the pause data generated by the Translog key stroke logger, which uses a

one-second default pause threshold, and also because this pause length has been used effectively in several previous translation keystroke logging studies. See Alves and Vale (2009) for a discussion of how long an interruption in the text production process should be in order to be considered a pause as well as their argument for the one-second threshold.

2.7 Filled pause

Filled pauses are interruptions in speech that are filled with utterances such as "ahh," "uhm," "err," "uhh," and "hmm." For further discussion, see Goldman-Eisler (1961), Rochester (1973), and Shriberg (1994). Filled pauses are considered a type of pause because they are thought, like unfilled pauses, to delay speech production to, for instance, allow time for accessing the next word to be uttered, or to plan a syntactic shift (Beattie and Bradbury 1979). In the following example, the pause filled with the *uhm* disfluency is 1.16 seconds long.

(1) Santa Fe de Bogota also offers *uhm* everything that characterizes a big city.

2.8 Repetition

A repetition occurs when the speaker repeats part of the utterance (Fox Tree 1995). The repeated portions must be identical (or nearly so; we included incomplete words in repetitions). Repetitions may represent hesitation due to incomplete planning (Clark and Clark 1977). In the following example, the word *dedicated* is repeated three times, with the middle repetition incomplete.

(2) … to which he had *dedicated ded dedicated* the last 12 years of his life exclusively.

2.9 Repair / revision

A repair occurs when an utterance is repeated, but with an alteration to one or more constituents (for instance by replacement, omission, or insertion of words), and the other constituents generally preserved (Shriberg 1994). Using Shriberg's terminology, the initial segment is called the *reparandum* or RM, and the repairing segment the *reparans* or RS. In the example below, the RM *significant section* is revised to the RS *significant number*. We have not used the term *restart,* which is also frequent in the literature as a near synonym for revision or repair, preferring to classify all utterance modifications as repairs.

(3) Spanish language is considered a symbol of national identity for a *significant section significant number* of the Venezuelan population.

2.10 Analytic protocol

To code the occurrence of speech disfluencies, we adopted an *annotated protocol*. The rationale for the protocol is that we believe an appropriately detailed transcript of sight translation performance, produced by annotating both errors and speech disfluencies, can act as a kind of *verbal protocol*. Much like a concurrent (think-aloud) or retrospective protocol, it can provide significant information to translation process researchers interested in language mediation with verbal output such as sight translation, verbal gisting (written to spoken summary translation), and, of course, interpreting. When used in studies in combination with written translation output and compared to detailed analyses of key stroke logs, annotated protocols might function to highlight significant differences in information processing associated with the modality of the language mediation.

To create the annotated transcripts, a method was adopted that allowed speech and translation quality data to be described in a single data representation. The method enables the description of discrete *speech performance properties* and *speech performance events*. Performance properties are location, timestamp, and duration. Performance events include a whole set of speech phenomena such as pauses, verbal articulations (either self-addressed or experimenter-addressed) that are not part of the translation, repetition of source text segments, revision of target text productions, repeating of target text words or segments, incomplete words, and other speech artifacts.

For each of the participants in our original experiment, we had set Tobii's Clearview software to produce a video transcript of the sight translation in the form of an AVI file. We then used an audio processing program called Sound Forge© to strip the audio track from the video file and save it as a WAV file. The audio processing software was then used to display the audio of the participant's sight translation as a waveform. Using the program's display window at a 1:256 zoom ratio, we were able to identify and locate performance events such as speech disfluencies and pauses. Using the program's timeline and editing features, we were able to provide timestamps and measures of durations for such events. In Figure 2 below, for example, we have selected and measured a pause beginning at timestamp 2.250 and with a length of 0.288 milliseconds.

From listening to the audio file, a transcript could be prepared and then annotated with time and duration information using the timeline information provided by the software. The annotation uses a set of arbitrary conventions for indicating time stamps (pair of curly braces, {}), unfilled pauses (pair of slashes, //), filled pauses (pair of square brackets, []), incomplete words (an exclamation point, !), repetitions (pair of ampersands, &&), revisions (pair of pounds signs, ££), and the targets of quality evaluations provided by an evaluator/rater (pair of asterisks, **). The annotation markers *enclose* the relevant data. Thus, we can see below in example (5) that *&indi! Indigenous&* encloses a repetition. Any desired annotation convention could be used, but the scheme shown here could easily be converted to extensible markup language (XML) format via a regular expression parser.

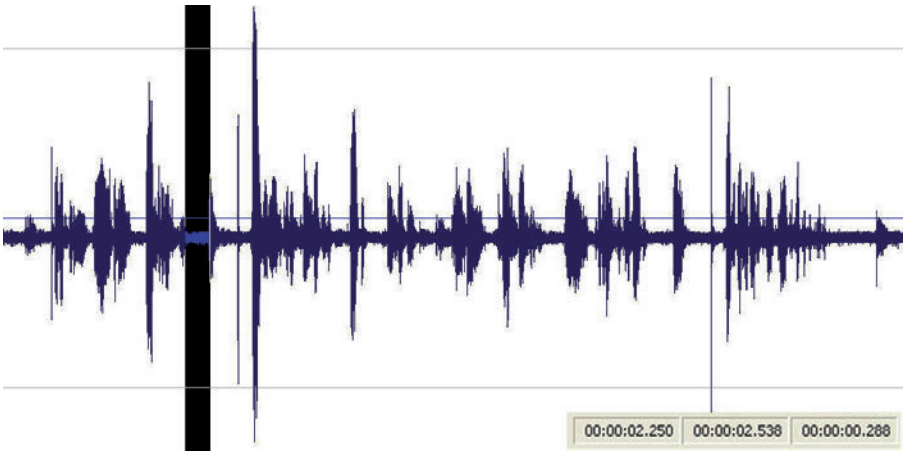


Figure 2. Segment of audio transcript in waveform view

(4) Transcript: Also, too, there are *oral traditions*, *uhm tsk* valuable oral traditions, in other languages like, for example, the *indi!* indigenous languages, these traditions are not *uhm* part of the understanding of all of the inhabitants of the country.

(5) Annotated Protocol: *Also, too*, there are *#oral traditions, [UHM TSK = 1.717] {1:45.290} valuable oral traditions#*, in other languages like, for example, the &indi! indigenous& languages, these traditions are not [UHM TSK = 2.077] {1:57.609} *part of the understanding* of all of the inhabitants of the country.

The protocol shown in example (5) above documents the (1) location and duration of all within-sentence unfilled pauses of 1 second or greater in length (major pause), (2) time stamps for the beginning and ending of sentences as well as for all non-propositional disfluencies and major pauses, (3) duration of all between-sentence pauses regardless of length, (4) location and duration of all filled pause disfluencies regardless of length, (5) location of incomplete words, repetitions, or revisions, and (6) scope of the object of a rater's evaluation. The time stamps are particularly useful because they make it possible to cross-reference certain performance events, such as major pauses and disfluencies, with eye-tracking events, such as fixations and regression, whose timestamps have been captured by the eye-tracking software.

Transcripts were prepared in a three-step process. A preliminary verbatim transcript, numbered by sentence, was prepared by one of the researchers directly from the audio file. The transcript included all sounds and noises, but did not include time stamps, pause indicators, or other annotations. This transcript was reviewed by a member of the research team who was a native Spanish speaker and a translation instructor. This reviewer provided quality evaluations, explicitly indicating the linguistic scope of the evaluation in the transcript. There were no additional reviewers.

The transcript was then annotated by another researcher, who coded the disfluencies from a codebook prepared by the research team and indicated timestamps and durations. Transcripts were not annotated by a second annotator. The protocol described here is preliminary, and any future applications of the protocol would need to involve multiple reviewers and annotators, a method to address consistency of quality evaluation and annotation, and reporting of inter-reviewer/annotator reliability measures.

3. Results

The annotation protocol was applied to twenty transcripts representing *A-Text* and *B-Text* versions of 4 different texts. Each of the ten participants completed an *A-Text* version of one of the four texts, and a *B-Text* version of another. The number of source text words and sentences processed is given in Table 1. The source texts were used as the basis for calculating per word and per sentence disfluency rates.

3.1 Frequency and duration of disfluencies

The frequencies of the four speech disfluencies we studied are given in Table 2. The results indicate that filled pauses are by far the most frequent speech disfluency found, followed by repairs and unfilled pauses. The rates of disfluency per source text word and sentence processed (number of words and sentences processed in the experiment by each participant) are also given. Fox Tree (1995) estimated that the disfluency rate in spontaneous speech is about 6 events per 100 words uttered. From this ratio we might expect, if sight translation were equivalent to normal speech, about 200 disfluencies in the 3,347 words uttered. We can see from Table 2 that overall disfluency rates from the experimental transcripts are over three times higher, with a disfluency

Table 1. Source corpus

Source text	Words	Sentences	Participants	Total words by participants	Total sentences by participants
1A	167	10	3	501	30
1B	167	10	2	334	20
2A	171	9	2	342	18
2B	172	10	3	516	30
3A	170	7	3	510	21
3B	170	7	2	340	14
4A	159	7	2	318	14
4B	162	9	3	486	27
TOTAL	1338	69	20	3347	174

Table 2. Frequency and rate of disfluency (DF) types in sight translation transcripts

Disfluency type	Frequency	DF count	Rate ST word	Rate ST sentence
Unfilled pause	17.38%	118	0.04	0.68
Filled pause	47.28%	321	0.10	1.84
Repetition	13.25%	90	0.03	0.52
Repair (restart)	22.09%	150	0.04	0.86
Total	100%	679	0.20	3.90

Table 3. Unfilled (UFP), filled (FP) and inter-sentential pause (ISP) durations

Text	UFP duration	Count	Average	FP duration	Count	Average	ISP duration	Average
ID	31.352	14	2.2394286	223.414	81	2.758198	129.931	2.887356
1	44.195	29	1.5239655	113.788	59	1.92861	62.658	1.457163
2	51.253	32	1.6016563	139.26	74	1.881892	47.628	1.5876
3	90.194	43	2.0975349	302.72	107	2.829159	96.362	2.676722
4	216.994	118	1.8389322	779.182	321	2.427358	336.579	2.185578

rate of 679. If speech disfluencies, as has been proposed in the literature, are indicative of cognitive effort, then relative to spontaneous speech sight translation appears significantly more effortful.

Table 3 gives the relative duration and count values for unfilled, filled, and inter-sentential pauses in the four texts we used. To simplify presentation and because the purpose is to compare frequency of pause types, we collapsed A and B version pause values into a single total. We can see that filled pauses not only occur more frequently than unfilled pauses, but are on average 31% longer. Inter-sentential pauses are only slightly shorter than filled pauses.

3.2 Disfluency evidence for visual interference

Results using eye tracking metrics showed the predicted syntactic effect in sight translations for *A-Texts*, when the non-complex AOI occurred in the first paragraph and the complex AOI in the second. The differences were statistically significant and showed that paragraphs containing complex AOIs were more effortful to process than paragraphs containing non-complex AOIs. The effect was not found in sight translation *B-Texts*, when the complex AOI occurred in the first paragraph and the non-complex in the second. In trying to explain this result, a more detailed statistical analysis was conducted, where first paragraph AOIs from *A-Texts* were compared with first paragraphs AOIs from *B-Texts*. This comparison showed the syntactic effect.

To explain the *B-Text* results in the sight translations, we proposed that the syntactic effect was likely masked in that condition by *visual interference*, a phenomenon

Table 4. Speech disfluencies before and after text midpoint

Midpoint	Measure	Text 1	Text 2	Text 3	Text 4	Total
Before	DF count	59	65	73	81	278
	UFP/FP duration	93.747	74.85	83.166	124.484	376.247
After	DF count	76	83	93	144	396
	UFP/FP duration	161.019	83.133	107.347	268.43	619.929

previously identified in sight translation by Gile (1997) and Agrifoglio (2004). Visual interference is a disruption induced by the continued presence of the written source text during the course of the sight translation. It appears to be temporally related, that is, the effect of the interference seems to increase as the course of the sight translation progresses. This created a situation where all second paragraphs were apparently more effortful to process than first paragraphs. There was no evidence of visual interference in the written translation task, indicating that the phenomenon was specific to sight translation. For a more detailed discussion of the eye tracking metrics and a more detailed discussion of the evidence for visual interference, see Shreve, Lacruz and Angelone (2010).

Given the findings in our initial paper, we wanted to see whether the disfluency findings might corroborate the conjecture of visual interference as a corollary of sight translation. For each of the participant protocols, we determined the midpoint of the source text and then calculated the frequency and duration of disfluencies before and after the midpoint. The assumption would be that if the visual interference apparent in the eye tracking metrics is related to increased cognitive effort as the translation progresses, then other phenomena associated with effort, such as speech disfluencies, might also indicate the effect of the interference. We should then see more disfluency after the midpoint of the sight translation than before it.

As can be seen in Table 4, the trend strongly supports the assumption of some kind of interference effect, as the second half of each of the four texts contains more disfluencies than the first half, and more time is spent in filled and unfilled pauses. The before and after midpoint values for the A and B versions of each text are added together because the visual interference hypothesis would predict that *regardless* of whether the first half of the text contained a complex or a non-complex AOI, the second half of the text would have more and longer disfluencies. The disfluency data corresponds very well with the trend apparent in the eye tracking data where the second halves of texts shows increased measures of cognitive effort (number of fixations, fixation duration, and number and duration of regressions).

3.3 Disfluency evidence for the effect of the syntactic manipulation

Since the study also involved a syntactic manipulation, we wanted to determine whether the disfluency patterns were analogous to the eye tracking patterns associated

Table 5. Disfluency frequency and duration aggregated by area of interest type

AOI	Measure	Text 1	Text 2	Text 3	Text 4	Total
Non complex	DF count	18	35	45	54	152
	UFP/FP duration	31.29	37.308	55.556	94.377	218.531
Complex	DF count	22	49	39	60	170
	UFP/FP duration	30.272	41.051	45.123	113.251	229.697

with the syntactically complex and non-complex areas of interest. The complex areas of interest are the most syntactically complex sentences in their respective texts, and if complex syntactic operations require cognitive effort, we would expect disfluencies to accompany them. Table 5 shows the frequency of disfluency events and the duration of filled and unfilled pauses when events occurring in non-complex and complex areas of interest from *A-Text* and *B-Text* versions are aggregated and then compared by AOI type. The table shows some difference between the complex and non-complex areas of interest, with complex areas showing 11% more disfluent events and lasting 5% longer. However, if the conjecture of visual interference is true, then some of the syntactic effect is being masked by interference as the participant approaches and passes the midpoint of the text.

If we take a more detailed look at the distribution of events and durations, a different picture emerges. Table 6 breaks down event frequencies and durations by the location of the area of interest in the text. There are three cases for comparison: (1) when we compare the non-complex first paragraph AOI of *A-Texts* with the complex first paragraph AOI of *B-Texts*, there appears to be a strong effect, with events increasing 34% (82/61) and duration 22% (96.233/78.855); (2) similarly, when we compare the first paragraph non-complex AOI of *A-Texts* with the complex AOI in their second paragraphs, we also seem to see a strong effect, with events increasing 44% (88/61) and duration 70% (133.554/78.855); but (3) the effect seems to disappear when the non-complex AOI occurs in the second paragraph of *B-Texts*. In fact, the non-complex AOI in this case is about 11% more eventful (91/82) and 45% longer (139.676/96.233) than the complex AOI in the first paragraph. These disfluency results track amazingly well with the eye tracking data where the exact same trends were noted. The results would seem to corroborate the "masking" effect of visual interference. Of course, as with the eye tracking metrics, the results also raise the question of how much of the effect we see in the second case is contributed by visual interference and how much by the cognitive effort associated with the complex syntax.

3.4 Individual protocols as evidence of cognitive processing

In addition to providing the aggregate data reported above, the application of our analytic protocol made it possible for us to mark the location and important characteristics of individual disfluencies in discrete sentences in the corpus. This capability

Table 6. Disfluency frequency and duration aggregated by location of AOI

AOI position	Type	Measure	Text 1	Text 2	Text 3	Text 4	Total
			Case 1				
			1A	2A	3A	4A	
1st Paragraph (A-Text)	NC	Count	6	15	27	13	61
		Duration	8.786	12.369	33.647	24.053	78.855
			1B	2B	3B	4B	
1st Paragraph (B-Text)	C	Count	10	26	16	30	82
		Duration	17.712	24.773	15.191	38.557	96.233
			Case 2				
			1A	2A	3A	4A	
1st Paragraph (A-Text)	NC	Count	6	15	27	13	61
		Duration	8.786	12.369	33.647	24.053	78.855
			1A	2A	3A	4A	
2nd Paragraph (A-Text)	C	Count	12	23	23	30	88
		Duration	12.65	16.278	29.932	74.694	133.554
			Case 3				
			1B	2B	3B	4B	
1st Paragraph (B-Text)	C	Count	10	26	16	30	82
		Duration	17.712	24.773	15.191	38.557	96.233
			1B	2B	3B	4B	
2nd Paragraph (B-Text)	NC	Count	12	20	18	41	91
		Duration	22.504	24.939	21.909	70.324	139.676

is extremely important methodologically. The use of a detailed protocol allows the researcher to distinguish between information provided by the *translation product* and information provided by the *translation performance*. The distinction between product and performance is implied in other data acquisition methods, as for instance key stroke logging. Key stroke logging programs such as Translog also use an annotation scheme to capture performance data. We argue that our performance protocols can perform an analogous function for oral output. Like key logs, they allow us to pay attention to discrete aspects of the way the translation unfolds in real time and not just to what results. The ability to analyze performance in a rigorous way is extremely important, because it is through the analysis of performance that the most complete evidence of underlying information processing can be gathered.

To illustrate our point, let us look at the examples below. The Spanish source segment shown in (6) is the stimulus for the sight translator to render the target production in (7). This production represents, not what we *hear*, but more likely what we *attend to* and *remember* as listeners. The translation production is analogous to a written translation product, where all of the hesitations, false starts, pauses, and revisions bound up with that activity are ephemeral and, unless captured by a keystroke logger,

are lost in the final version offered up to a reader. In speech performance, listeners most often gloss all but the most marked disfluencies. In fact, speakers themselves most often do not note or recall the production of disfluencies. Example (8) represents what is actually performed and heard. There is valuable evidence in the acoustic performance data that can be exploited in translation process research. Example (9) shows a marked-up version of the performance with disfluency events annotated and problem areas marked by a rater (in bold between asterisks).

(6) SOURCE SEGMENT: Si bien hay tradiciones orales valiosas en otras lenguas como, por ejemplo, en las lenguas indígenas, estas tradiciones no son del dominio de todos los habitantes del país.

(7) TRANSLATION PRODUCTION: Also, too, there are valuable oral traditions in other languages like, for example, the indigenous languages; these traditions are not part of the understanding of all of the inhabitants of the country.

(8) TRANSLATION PERFORMANCE: Also, too, there are *oral traditions*, *uhm tsk* valuable oral traditions, in other languages like, for example, the *indi!* indigenous languages, these traditions are not *uhm* part of the understanding of all of the inhabitants of the country.

(9) TRANSLATION PROTOCOL: {1.40.078} *Also, too*, there are *#oral traditions, [UHMTSK=1.95] {1:45.290} valuable oral traditions#*, in other languages like, for example, the &indi! indigenous& languages, these traditions are not [UHMTSK=2.077] {1:57.609} *part of the understanding* of all of the inhabitants of the country.

So, what can we discover from the annotated protocol above? First, we see that there are two fairly long filled pauses of around two seconds. The first of these follows the reparandum *oral traditions* and precedes the reparans *valuable oral traditions*. This segment has been marked by the translation rater. The second precedes a short phrase *part of the understanding* that has also been marked by the translation rater.

These long pauses, if they are indicative of cognitive effort (as the disfluency literature indicates) should be associated with groups of eye fixations. In fact, we can see from the "hot spot" image produced by the Clearview software (see Figure 3) that there is clearly a group of fixations centered near the first occurrence of the Spanish word *tradiciones* in line 6. The hotspots, which appear as lighter colored areas, indicate areas where there is a larger concentration of eye fixations. Even more apparent are the major fixations clustered around the phrase *estas tradiciones no son del dominio de todos* (line 7), the site of the two-second filled pause. Even the repetition of the word indigenous is associated with minor fixations at *las lenguas indígenas* in line 7.

It is not the goal of this paper to discuss in great detail the specific correlations between the eye tracking and disfluency data, since our initial design was not set up to provide evidence of any such linkage. Nevertheless, the apparent correlation of disfluency with eye tracking data in our small corpus of results is quite suggestive.

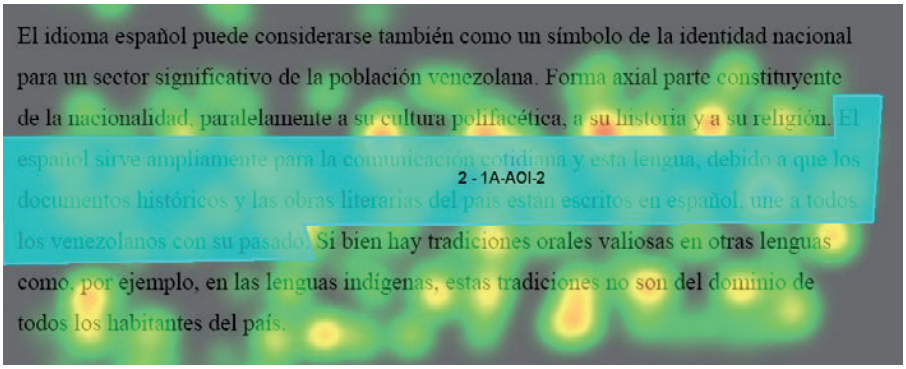


Figure 3. Eye tracking plot of fixation hot spots

Methodologically, it is important to emphasize that eye phenomena, such as fixations and regressions, are part of the translation performance. If it is possible to use a richly-annotated protocol to associate one category of performance data with another, then we have enriched the information available to interpret the progress and character of underlying cognitive processes.

What else can we determine from the disfluency data? The apparent correlation of the filled pause locations with eye fixations indicates that some disruption to cognitive processing has occurred in the sentence. The location of the pause in the sentence is potentially quite significant. Many speech disfluencies are prospective (Bell, Jurafsky, Fosler-Lussier, Girand and Gildea 1999), signaling upcoming syntactic or lexical difficulties. In example (9) above, for instance, we can see that the participant has encountered a syntactic production difficulty. The translator produces the segment *oral traditions* very quickly (signaling most likely a failure to scan ahead), but then there is a pause. The translator finally notes that the source phrase that cued the first utterance (perhaps quite automatically) *tradiciones orales* is actually *tradiciones orales valiosas* in its entirety. The translator must then take account of differences in Spanish and English word order and repair the utterance, inserting the adjective *valuable* to produce the repaired utterance: *valuable oral traditions*. The disfluencies (an unfilled pause and a repair) are associated with the identification of the mandatory syntactic transposition and its planning.

Our analysis of the location and nature of disfluencies, when combined with an evaluator's comments, allowed us to associate the disfluencies with probable sources of processing difficulty. The great majority of disfluencies noted in the corpus of 20 transcripts were syntactically and lexically related. We did note, however, that some of the disfluencies seemed to signal production problems related to choosing a translation strategy.

3.5 Syntactic transfer disfluencies

Examples of syntactic planning and production marked by disfluencies abound in our small experimental corpus. In the same text we examined earlier, the translator makes several more syntactic repairs. These are not marked by filled pauses, but we can clearly see evidence of the recognition of a production problem and its resolution with a syntactic shift.

(10) … paralelamente a su cultura polifacética …

(11) … parallel to its *#culture, multi-faceted culture#* …

(12) … debido a que los documentos históricos …

(13) … according to *#documents, historical documents#*…

These changes are not idiosyncratic, for we can see another participant making a syntactic repair in the same location:

(14) Si bien hay tradiciones orales valiosas …

(15) …there are *#oral traditions, valuable oral traditions#* in other languages …

And with yet another participant we have the same transposition, this time accompanied by an unfilled pause, probably associated with lexical access to the equivalent of *tradiciones* followed by a filled pause, here associated with the repair of *traditional* to *traditions* and the syntactic transposition.

(16) There are even* /3.320/ {4:10.705} *#traditional /UHH=1.555/ {4:13.004} oral traditions valued oral traditions#*

With yet another participant, we see a complex (and mostly failed) syntactic repair, where *valiosas* is rendered originally as *good*, then the adjective *oral* is added, and finally there is a repair of *good* to *valuable*.

(17) there are *#good traditions, oral traditions# valuable*

This text was translated by five participants, and what is very interesting is that a production problem associated with contrastive syntax is marked by repair disfluencies in four out of the five translations.

The translation of more complex syntactic constructions, as one would expect, is also marked by disfluencies. Here, an unfilled pause of 1.822 seconds and a short repetition precede the transfer of a dense syntactic construction. After the pause, the translator renders *historical and touristic site, places, interest*. Pausing again for almost a second, a repair is attempted, *historical and cultur and tour touristic touristic places of interest*. An optimal solution is never rendered. Although the individual words are relatively simple and appear to pose no problem, the translator becomes confused because the first word *sitios* is a noun and she assumes the second word *interés* to be

an adjective accompanying it. But the word is actually another noun qualified by the two subsequent adjectives *histórico* and *turístico*. So the construction in Spanish is [noun + (noun + adjective + adjective)], which should translate as *places of historical and touristic interest*.

(18) Además de los sitios de interés histórico y turístico, Santa Fe de Bogotá también ofrece todo lo que caracteriza a una gran metrópolis.

(19) *#Besides &the /1.822/ {1:40.728} the& historical and touristic #site, places#, interest /UHM=0.951/ {1:46.068} besides the historical and cultur! and &tour! touristic touristic& {1:49.58} places of interest #*, Santa Fe de Bogota also offers … everything that characterizes a big city …

Many sentences show multiple syntactic effects. In our next example, we see a 1.637-second unfilled pause followed by a syntactic transposition to accommodate the adjective *pre-Columbian*. This is followed by a 1.021-second unfilled pause marking a problem with the production of a relative pronoun. The source text has an embedded relative clause starting with *cuyo*. The word can be translated as *whose* or *which* depending on the context. The translator starts with *who* and then repairs to *which* (as an incomplete word) and finally decides to ignore the relative pronoun and omit it (location marked by OMIT).

(20) … el cual alberga una colección de más de 36.000 piezas de orfebrería precolombina, algunas de las cuales son extraordinarias, cuyo género es único en el mundo.

(21) … it's home to a collection of more than 36,000 /1.637/ {1:25.379} pieces of *#goldwork, pre-Columbian goldwork#*, some of which are quite extraordinary and /1.021/ {1:32.461} *#who whi!* OMIT and# are unique throughout the world.

3.6 Lexical transfer disfluencies

Production difficulties due to contrastive or difficult syntax are not the only sources of disfluency in our transcripts. As indicated by the disfluency literature, problems associated with lexical access and lexical choice may also be associated with disfluencies. In translation tasks, we would certainly expect lexical transfer to be a significant source of cognitive effort, particularly with low frequency source language words, abstract words, words without direct equivalents in the target language, or words with multiple equivalents (e.g. where lexical choice is a critical issue).

In the following example, there is a problem with finding appropriate equivalents. The problem words are *muestra*, *huellas*, and *vestigios*. An entire complex of disfluencies accompanies the difficulty associated with access and retrieval of equivalents for these words from bilingual memory. There is a massive filled pause of over 11 seconds prior to the production, which is supposed to render these three words. Unable to

retrieve equivalents, the participant stalls by repeating the words *we still find*. This is followed by yet another repetition, of the article *a*. The equivalent *samples* for *muestra* is ventured and is apparently not deemed a satisfactory rendering, because another pause of 3.482 seconds ensues, followed by an attempt at a repair with the word *examples*. Equivalents for the two other Spanish words (*huellas* and *vestigios*) are never produced.

(22) Allí aún nos encontramos con una muestra de huellas y vestigios de un pasado glorioso.

(23) There … &we still find /UHM-TSK-SUBVOCAL]=11.435/ {0:34.806} *we still find& &a a& #sample of /UHM=3.482/ {0:42.120} examples* of# OMIT the glorious past.

In the following example, another participant translating the same text runs into similar difficulties in the same location. Here the translator encounters a similar problem with retrieving an equivalent for *muestra*. She has a long pause of 5.607 seconds, and then attempts the word *demonstration* as an equivalent. Apparently satisfied with the solution (even though the direct equivalent *sample* would have been more appropriate), the translator proceeds directly into a difficulty with *huellas*. This is associated with an interesting filled pause, where the participant repeats the Spanish word and then says sub-vocally, "I know that word." The translator then omits the word *huellas* from the translation and proceeds to find an equivalent for *vestigios* by using the cognate *vestiges*.

(24) There *even* we find with /UHM=5.607/ {0:27.736} *demonstration* of the [SPAN: huellas I KNOW THAT WORD]=1.369 {0:31.056} OMIT even vestiges of a glorious past.

The corpus of transcripts contains a whole host of interesting examples of disfluencies associated with lexical transfer. From the same participant later in the text, we see an attempt to find an equivalent for *aridez*. This word has a direct English cognate, *aridity*. But instead of using the cognate, the translator creates a pseudo-cognate, a plausible English non-word *aridation*. The filled pause (where the translator laughs) most likely indicates uncertainty with the choice and may be evidence of what Angelone (2010) has called a *solution evaluation*.

(25) Al mirar la aridez de la tierra …

(26) Upon seeing the *#aridation like [LAUGH=0.336] {2:13.607} aridation#* of the land …

While many lexical transfer disfluencies in our transcripts appear to be associated with the search for equivalents, they are not the only lexically-related phenomena we observed. For instance, in the text below, we see quite another situation. Here the translator retrieves the proper equivalent for *microcosmos*, but then apparently struggles with the issue of whether to make the equivalent singular or plural. The participant

hesitates and then translates it correctly in the singular as a *type of microcosm*, but then apparently notes that the word in Spanish ends in an "s" and decides, marked by a 1.822-second unfilled pause, to put it in the plural as a *collection of microcosms*. Here the issue is not lexical equivalence, but the transfer of lexical features such as number.

(27) Esto convierte a Santa Fe de Bogotá en una especie de microcosmos.

(28) This has made Santa Fe de Bogota *one /ERR=0.603/ {0:30.859} #a type of microcosm or, /1.822/ {0:34.249} a collection of microcosms#*.

In this final example, the disfluencies seem to indicate a concern with register. While we found very few register-related disfluencies overall in our small corpus, they represent an interesting case and might be expected to occur more frequently in the performances of more expert translators (or those more skilled in oral output). Here the translator renders *found* for *encontro* first, but then repairs to the cognate *encountered*, which is of somewhat higher register. Interestingly enough, she goes back to *found* with a second repair, apparently trying to avoid using the cognate. Later in the sentence, she renders *ultimo reposo* as *ultimate rest*, which is acceptable, but then she repairs it with the higher register *ultimate repose*, although it is unclear why she did not use *repose* in the first instance. We discuss a possible strategic explanation for this apparent avoidance of cognates in the next section.

(29) Gaudí encontró el último reposo en el lugar donde había trabajado 43 años de su vida y al que había dedicado sus 12 últimos años en exclusiva.

(30) Gaudi /1.056/ {1:54.671} *#found, encountered*, his *found*# #his ultimate *rest, his ultimate repose#*, in the place where he had worked for 43 years of his life /1.137/ {2:05.817} and to which he had dedicated his last 12 years exclusively.

3.7 Translation strategy and disfluency

While the great majority of disfluencies were associated with the syntactic and lexical production problems we have noted above, there were a number which were not so easily classified into those two categories. We would like to propose that some disfluencies are associated with translation decision making and mark the translator's conscious deliberation of text production choices, e.g. translation strategy. For instance, the source text sentence given below is relatively simple and contains no problematic words or syntactic constructions. It could be rendered by a direct translation strategy quite successfully. An examination of the source text yields no obvious lexical or syntactic explanation for the long pauses we find in the oral performance. But the translator is obviously looking to recast the sentence. We have a 3.644-second filled pause followed by a relatively innocuous repair. Then we have a total of three pauses, two filled and one unfilled, that seem to be associated with attempting an unnecessary recast. It is as if the translator thought that a direct translation could not be the right

strategic choice, and decided to recast at all costs. As a result, the translator ends up with some hybrid production *is used as communication* where the source text says *is used in the communication*.

(31) El español también se usa en la comunicación del país con otras naciones.

(32) Spanish #is also used in /UHM-TSK-AHH=3.644/ {1:27.992} *#is used# /UHM=1.252/ {1:29.652} between /1.729/ {1:32.090} the country and other nations is used as communication* between /UHM=0.649/ {1:37.071} country and other nations#.

While the psycholinguistic and interpreting literature indicates that most speakers are unaware of or unable to recall the reasons for the disfluencies they produce, in situations of strategic disfluency, it might be fruitful to see whether retrospective protocols prompted by the replay of the screen recording and a transcript of the translation might shed some light on the probable causes of strategic disfluencies such as this one. These reports could be combined with (in our case) eye tracking data or other processing data. The efficacy of such an approach is argued in Englund and Tiselius (2009: 129): "retrospective data cannot be taken as sole evidence for cognitive processes or strategy use. However, it can yield interesting results when combined with other methods or as part of triangulation."

Another example of strategic disfluency emerged in the way some of the translators handled cognates. We saw two examples of this in the previous section with respect to changes in lexical register. In the following example, there is a 1.323-second pause preceding the production of *a symbol*. There is no explanation for this pause, since *símbolo* and *symbol* are direct cognates. The implication is that the translator recognizes the two words as cognates, but is assessing the wisdom of applying a translation strategy using the cognate. In this case, the strategy is applied. This first instance involving a cognate is followed by a very long 4.968-second pause, before *for a significant*, and then an unfilled pause of another 3.529 seconds before *section*. In this case, the source text word *sector* has a direct English cognate, but the translator seems to avoid using it, offering instead several possible quasi-synonyms (*section*, *number*).

(33) El idioma español puede considerarse también como un símbolo de la identidad nacional para un sector significativo de la población venezolana.

(34) *Spanish language* /1.021/ {1:42.295} is considered /UHH=1.323/ {1:45.571} *a symbol* of national identity /UHM=4.968/ {1:51.943} *for a #significant /3.529/ {1:57.516} section significant number#* /UHH=3.389/ {2:03.646} of the Venezuelan population.

This strategy of *cognate avoidance* was not unique to this translator and was observed in several other cases. In the transcript of another translator, we see a repetition that might signal some uncertainty about using the cognate *symbol*. We also have a half-second pause before *significant sector*, where the most likely explanation for the disfluency is cognate avoidance. In this case, the translator decides to use the cognate.

(35) The Spanish language can also be considered as a *&sy! symbol&* of the national identity /UHM=0.545/ {1:4.806} for a significant sector of the Venezuelan population.

Shlesinger and Malkiel (2005) report some interesting results on the translation of cognates by professionals, comparing interpreting and written translation. They found that non-cognate synonyms are more prevalent in translation than in interpreting and that performance on false cognates was superior in translation. This would seem to indicate that the increased time allowed in written translation for problem resolution allows the translator to consider other options than direct translation of the cognate and to factor in the possibility of *faux ami*. We might expect similar results if we compare our sight and written translations. In addition, it might be a fruitful avenue of research to see if interpreters exhibit any signs of cognate avoidance. Avoidance could be ascribed to difference in modality and the associated time constraints, but it might also be due to the training of the participants, with student translators perhaps more likely to mistrust direct cognate translation than professionals.

Another example of strategic disfluency can be seen in the rendering of idioms. Here the idiom *hacia tiempo*, most usually translated as *time ago* does not fit the sentence context. The translator pauses with an unfilled pause for 1.462 seconds, considers the strategic alternatives, and then produces a very good rendition, *in his time.* Interestingly, there is an unfilled pause of similar duration (1.567) following the translated phrase, which does not appear to be associated with the following phrase. This pause is of interest not just because it shows a case of disfluencies being associated with idiom translation, but because the pause after the rendition of the idiom may be a marker for what Angelone (2010) has called solution evaluation, the third stage of a three-stage translation problem solving process that includes problem recognition and solution proposal. In the solution evaluation stage, a translator assesses a proposed solution and either accepts it, or proposes another. In fact, it is possible to see many of the disfluencies in this corpus of transcripts fitting into Angelone's problem solving framework.

(36) Gaudí se había convertido desde hacia tiempo en un héroe popular.

(37) Gaudi had /1.462/ {1:28.108} *become in his time* /1.567/ {1:31.231} a popular hero.

In the following example, we see a very interesting phenomenon, indicative of translation-strategic operations occurring over the course of several sentences. In sentence (39), we see that *un cortejo fúnebre* is rendered as *a short funeral.* This is most likely because the translator misinterprets *cortejo* (procession) as "*corto*" (short). Perhaps the translator was attributing some function (informality?) to the ending *ejo* as the result of interference from the orthographic similarity of the forms. However, the word *cortejo* as a morphological derivation of *corto* does not exist in Spanish. There is no disfluency associated with the error, so it is apparently made and not remarked. Four sentences later, the word *cortejo* occurs again. Here a repair, *procession, funeral*

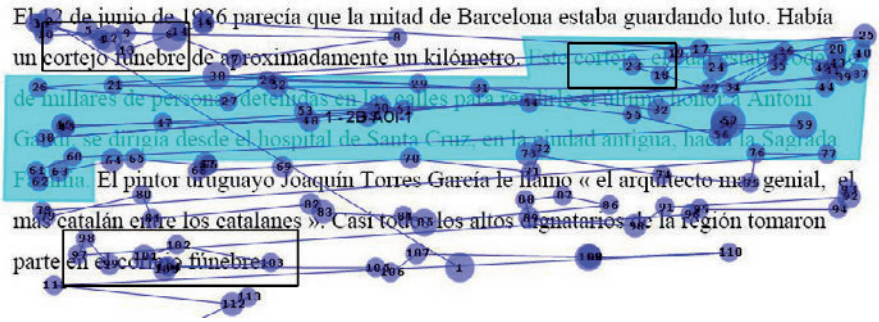


Figure 4. Gaze plot indicating fixations in paragraph containing *cortejo*

procession, occurs. Note, however, that the mistake made four sentences prior does not recur. This seems to be an example of what we might call a *delayed correction,* postponed for several sentences because the recognition of the error occurred after an opportunity to correct it had already passed. The correct production was never uttered, but a correction was nevertheless made and stored in memory for later implementation. Such patterns are intriguing because they may be evidence of some ongoing mental representation of the translation held in memory. They possibly indicate that certain disfluencies, in this case a repair, may mark strategic operations delayed in their execution over the scope of several sentences.

(38) Había un cortejo fúnebre de aproximadamente un kilómetro.

(39) *There had been* *a short funeral* of approximately one kilometer.

(40) Casi todos los altos dignatarios de la región tomaron parte en el cortejo fúnebre.

(41) Almost all of the higher figures of the region took part in this *#procession, funeral procession#*.

In the example above, disfluency marks the site of the delayed repair. Of course, the repair may take place without being marked by disfluency, as in the following example from a different participant translating the same text. Here the translator repeats the mistranslation of *cortejo* in two adjacent sentences, but then three sentences later corrects the error. The lack of a disfluency marker serves as a reminder that disfluencies are only a possible indicator of cognitive effort and production problems. They do not mark every site of production difficulty. That is why we need to interpret phenomena such as repairs in the light of triangulated data. If we look at the eye data, a gaze plot (see Figure 3) for the paragraph shows us that the neighborhood of *cortejo funebre* in the second line (indicated by a rectangle) is fixated about seven times on its first occurrence. *Cortejo* is fixated only once, perhaps twice, later in the second line. However, *en cortejo funebre* is fixated about eight times on its third occurrence in the last

line. This high number of fixations corroborates an interpretation of a delayed repair strategy in the absence of a disfluency marker.

(42) There was a *short* funeral march &of, of& approximately one kilometer.

(43) This *short* march, #which *was /1.695/ {0:34.098} surrounded …

(44) Almost all of the high dignitaries of the region took part *in the funeral march*.

4. Discussion

In general, the disfluency patterns corroborate the findings of the eye tracking results reported in Shreve, Lacruz and Angelone (2010) with respect to the syntactic manipulation. They also indicate, as did the eye data, that visual interference is an important cognitive factor in sight translation. The disfluency data do not help us disentangle the complex relationship between increases in effort due to visual interference as the sight translation progresses and effort due to other factors, such as complex syntax. We hope new experimental designs may be developed that can account for the effects of visual interference so that it is possible to conduct experiments that measure the real effects not only of syntactic manipulations, but lexical, propositional, or textual ones as well.

Perhaps the greatest contribution of the disfluency analysis is a methodological one. The congruence between the eye tracking results and the disfluency results shows that disfluency analysis can be used effectively in combination with methods like eye tracking to triangulate experimental results. Disfluency analysis via an annotated protocol also seems to provide an important methodology for exploiting the audio component of screen recordings for extracting useful data in process research.

Speech disfluencies are components of natural speech that are viewed as having little or no semantic value or communicative function in the utterances where they occur. There is some debate over this definition, with several researchers claiming at least some interpersonal communicative function for disfluencies like filled pauses (Clark and Fox Tree 2002). Clark (2002), for instance, argues that although filled pauses do not add to the propositional content of the primary message, they are part of a *collateral message* where the speaker is intentionally providing information about the message (e.g. levels of uncertainty) to an interlocutor. Corley and Stewart (2008) argue that there is scant evidence for the intentionality of disfluencies.

In the case of sight translation, the probability that pauses or other disfluencies serve any communicative function is very low unless one wants to make the case that the translator has adopted the listening researcher as a proxy interlocutor. The frequency and placement of disfluencies in the sight translation data, as well as their apparent association with syntactic, lexical, and translation-strategic problems, argues for a primary association of the disfluencies with the cognitive and linguistic processes of the oral translation itself. The primary cause of the disfluencies, in consonance

with the psycholinguistic literature, would seem to arise from the cognitive effort associated with specific production problems resulting from the translation task: lexical transfer, syntactic shifts, and translation-strategic shifts.

This finding is partially consonant with the findings of some studies of disfluency in simultaneous interpreting. In interpreting, there is some influence from the speaker's pattern of disfluencies on the interpreter's pattern. However, Tissi (2000: 121), although she is hesitant to commit to a systematic difference in disfluency distribution between source and target texts in interpreting, notes a number of interesting differences in disfluency distributions between oral source texts and oral target texts, pointing, for instance, specifically to the length of certain pauses where she remarks that "unlike the original texts, TTs show a remarkable incidence of pauses from 2.5 to 5 and from 5 seconds up, which do not even exist in STs." If we assume that long pauses are associated with the increased effort of the language mediation task, then these long pauses are not unexpected and we see similar long pauses in the sight translation data from our corpus.

Bakti's (2009) study of disfluencies in simultaneous interpreting differentiates between what she calls *error-type disfluencies* (e.g. grammatical errors, contamination, false word activation, tip of the tongue phenomena, change, word ordering problems and slips of the tongue) and *hesitation* or *uncertainty-type disfluencies* (e.g. fillers, repetition, restarts, and pauses within the word). Though her study focuses primarily on error-type disfluencies, it does report an interesting result on the rate of restarts (repairs) in simultaneous interpreting. In both the output of trainee and expert interpreters, she found that restarts (repairs and revisions in our analysis) were by far the most common disfluency type. For trainee interpreters, restarts accounted for 26.5% of disfluencies, and for more experienced interpreters, restarts accounted for 36.5%. Repairs were not the most common disfluency in our data, coming in a distant second to filled pauses. Thus, although disfluency results in interpreting can inform and to some extent support our interpretation of sight translation results, it is clear that the cognitive requirements of the two tasks are quite different and distinct, and may yield different disfluency results. That conjecture, of course, awaits more rigorous study.

Simultaneous interpreting is much more closely bound to the presence of interlocutors than is sight translation. The rate of delivery, prosody, intonation, and disfluency patterns of the original speaker will influence the oral performance of the interpreter. The requirements of those listening to the interpreter (in non-laboratory settings) will also presumably influence the oral performance, as for instance in prompting the interpreter to maintain acceptable rates of speech output. Once these aspects of the performance are factored in we assume that the transfer requirements of the language mediation account for the differences between speaker performances and interpreter performances reported by Tissi, Bakti, and others (see especially Pöchhacker 1995).

In sight translation there is no speaker driving the pace of language mediation, although the presence of a listener can, as with interpreting, influence the rate of speech. In the laboratory setting of our experiment there was no overt listener and no specific impetus to urge participants to support normal rates of speech (as a result,

most of our participants did not end up speaking at usual conversational or interpreting rates). This implies that the disfluency effects may be ascribed mostly to task characteristics of the sight translation: visual interference, differences between oral source and written source texts in terms of syntax, grammar, and register, and, most probably, the novelty of the sight translation task itself. In further studies of disfluency in sight translation, it might be fruitful to use interpreters as participants to see whether experience with oral production will affect the rate of filled pauses and repairs. Some of the disfluencies we noted could be due to the novelty of the task.

Many of the repairs we have reported here and illustrated by examples are interesting because we have shown that they are associated with disfluencies. The fact that sight translators encounter difficulties in lexical transfer or performing syntactic shifts is not so surprising. It is not surprising that they might attempt to repair their flawed output. We have argued that disfluency analysis, especially if triangulated with another form of data such as eye tracking, can contribute to our interpretation of how the recognition of these difficulties and their resolution through repairs play out cognitively. We have noted, for instance, how disfluencies can indicate how and when problem resolution strategies are applied. In future experiments, it might be useful to see if patterns of disfluencies can be mapped to underlying problem resolution sequences such as the problem recognition, solution proposal, and solution evaluation stages that have been put forward by Angelone (2010). It might be quite fruitful to interpret sequences of repairs in the context of solution proposal or evaluation and their associated pauses in light of problem recognition (preceding pauses) and solution evaluation (succeeding pauses). Problem-solving is also highly associated with metacognition, and it might be quite profitable to attempt to triangulate disfluency analysis with concurrent and/or retrospective verbal protocols to see if disfluencies are indicators of metacognition, especially monitoring and control.

Finally, we have attempted to show how disfluency analysis can help discover and document evidence of strategic decision making by sight translators. We saw, for instance, evidence for what we called delayed repairs. Such repairs are potentially of great interest because they point to a complex underlying cognitive sequence where an error is recognized after it is made, a solution is proposed and evaluated without immediate production, and then retained over a span of several sentences before it is applied. Of equal interest is the apparent application of a cognate avoidance strategy by several of our participants. What does cognate avoidance signify? The phenomenon points directly to concerns that are not associated directly with lexical transfer or syntactic shifting. It points to a translator's assessment of the likelihood that a strategy that is being considered for implementation is the correct or most efficacious one. Cognate avoidance is evidence that translators are concerned not just with the linguistic mechanics of target language production, but with the sufficiency of their knowledge of translation. They may pause, repeat, and repair because they are assessing what they know and don't know about a particular act of translation. In cases like this, disfluency analysis leads us directly to questions of translation expertise, which is, uhm, uhh, the real heart of the matter.

References

Agrifoglio, M. 2004. "Sight translation and interpreting: A comparative analysis of constraints and failures." *Interpreting* 6 (1): 43–67.

Alves, F. and Vale, D. 2009. "Probing the unit of translation in time: aspects of the design and development of a web application for storing, annotating, and querying translation process data." *Across Languages and Cultures* 10 (2): 251–273.

Angelone, E. 2010. "Uncertainty, uncertainty management and metacognitive problem solving in the translation task." In *Translation and Cognition,* G. Shreve and E. Angelone (eds), 17–40. Amsterdam/Philadelphia: John Benjamins.

Bakti, M. 2009. "Speech disfluencies in simultaneous interpretation." In *Translation and the (Trans)formation of Identities: Selected Papers of the CETRA Research Seminar in Translation Studies*, D. De Crom (ed.), 1–18. Leuven: CETRA.

Barr, D. 2001. "Trouble in mind: Paralinguistic indices of effort and uncertainty in communication." In *Oralité et gestualité: Interactions et comportements multimodaux dans la communication,* C. Cavé, I. Guaïtella and S. Santi (eds), 597–600. Paris: L'Harmattan.

Beattie, G. W. and Bradbury, R. J. 1979. "An experimental investigation of the modifiability of the temporal structure of spontaneous speech." *Journal of Psycholinguistic Research* 8 (2): 225–248.

Beattie, G. and Butterworth, B. 1979. "Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech." *Language and Speech* 22: 201–211.

Bell, A. Jurafsky, D. Fosler-Lussier, E., Girand, C. and Gildea, D. 1999. "Forms of English function words: Effects of disfluencies, turn position, age and sex, and predictability." In *Proceedings of the International Congress of Phonetic Sciences,* J. Esling (ed.), 395–398. Berkeley, Calif.: IPA.

Bortfeld, H., Leon, S., Bloom, J., Schober, M. and Brennan, S. 2001. "Disfluency rates in conversation, age effects, relationship, topic, role and gender." *Language and Speech* 44 (2): 123–147.

Clark, H. 2002. "Speaking in time." *Speech Communication* 36: 5–13.

Clark, H. H. and Clark, E. V. 1977. *Psychology and Language: An Introduction to Psycholinguistics.* New York: Harcourt Brace Jovanovich.

Clark, H. and Fox Tree, J. 2002. "Using 'uh' and 'um' in spontaneous speaking." *Cognition* 84: 73–111.

Clark, H. and Wasow, T. 1998. "Repeating words in spontaneous speech." *Cognitive Psychology* 37: 201–242.

Corley, M. and Stewart, O. 2008. "Hesitation disfluencies in spontaneous speech: The meaning of 'um.'" *Language and Linguistics Compass* 2/4: 589–602.

Englund Dimitrova, B. and Tiselius, E. 2009. "Exploring retrospection as a research method for studying the translation process and the interpreting process." In *Methodology, Technology and Innovation in Translation Process Research* [Copenhagen Studies in Language 38], F. Alves, S. Göpferich and I. Mees (eds), 109–134. Copenhagen: Samfundslitteratur.

Fox Tree, J. 1995. "The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech." *Journal of Memory and Language* 34: 709–738.

Fox Tree, J. and Clark, H. 1997. "Pronouncing 'the' as 'thee' to signal problems in speaking." *Cognition* 62: 151–167.

Gile, D. 1997. "Conference interpreting as a cognitive management problem." In *Cognitive Processes in Translation and Interpreting,* J. Danks, G. Shreve, S. Fountain and M. McBeath (eds), 196–214. Thousand Oaks, Calif.: Sage Press.

Goldman-Eisler, F. 1968. *Psycholinguistics: Experiments in Spontaneous Speech.* New York: Academic Press.

Gósy, M. 2007. "Disfluencies and self-monitoring." *Govor* 26: 91–110.

Kemper S., Crow A. and Kemtes, K. 2004. "Eye-fixation patterns of high- and low-span young and older adults: Down the garden path and back again." *Psychology and Aging* 19: 157–170.

Levelt, W. 1983. "Monitoring and self-repair in speech." *Cognition* 14: 41–104.

Merlo, S. and Mansur, L. 2004. "Descriptive discourse: Topic familiarity and disfluencies." *Journal of Communication Disorders* 37: 489–503.

Pöchhacker, F. 1995. "Slips and shifts in simultaneous interpreting." In *Topics in Interpreting Research,* J. Tommola (ed.), 73–90. Turku: The University of Turku, Centre for Translation and Interpreting.

Rayner, K. 1998. "Eye movements in reading and information processing: 20 years of research." *Psychological Bulletin* 124: 372–422.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R. and Clifton, C. 1989. "Eye movements and on-line language comprehension processes." *Language and Cognition Processes* 4 (3/4): 21–49.

Rochester, S. R. 1973. "The significance of pauses in spontaneous speech." *Journal of Psycholinguistic Research* 2(1): 51–81.

Schachter, S., Christenfeld, N., Ravina, B. and Bilous, F. 1991. "Speech disfluency and the structure of knowledge." *Journal of Personality and Social Psychology* 60: 362–367.

Schnadt, M. and Corley, M. 2006. "The influence of lexical, conceptual and planning based factors on disfluency production." In *Proceedings of the Twenty-eighth Meeting of the Cognitive Science Society*, R. Sun and N. Miyake (eds), 750–755. Mahwah, N.J.: Lawrence Erlbaum Associates.

Shlesinger, M. and Malkiel, B. 2005. "Comparing modalities: Cognates as a case in point." *Across Language and Cultures* 6 (2): 173–193.

Shreve G. M., Lacruz, I. and Angelone, E. 2010. "Cognitive effort, syntactic disruption, and visual interference in a sight translation task." In *Translation and Cognition*, G. Shreve and E. Angelone (eds), 63–84. Amsterdam/Philadelphia: John Benjamins.

Shriberg, E. 1994. "Preliminaries to a theory of speech disfluencies." (Doctoral dissertation, University of California, Berkeley.)

Tissi, B. 2000. "Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis." *The Interpreter's Newsletter* 10: 103–127.

Traxler, M. J., Morris, Robin K. and Seely, R. E. 2002. "Processing subject and object relative clauses: Evidence from eye movements." *Journal of Memory and Language* 47: 69–90.

Time lag in translation and interpreting

A methodological exploration

Šárka Timarová*, Barbara Dragsted# and Inge Gorm Hansen#
*Lessius University College / K.U. Leuven, Belgium
#Copenhagen Business School, Denmark

Time lag between the source text input and the interpreter's target text is known as ear-voice span or décalage. Recently, time lag has also been measured in written translation based on eye-tracking and key-logging (hence called eye-key span). Time lag provides insight into the temporal characteristics of simultaneity in interpreting, speed of translation and also into the cognitive load and cognitive processing involved in the translation/interpreting process. Ear-voice span/eye-key span (EVS/EKS) thus have the potential to become very valuable measures in translation/interpreting process-oriented research, but more needs to be known about their properties and the information this metric yields. Several exploratory analyses were carried out to compare methods used in previous research and to address the questions of EVS/EKS variability among participants, stability of EVS/EKS over the course of an individual's performance, and variation in EVS/EKS at different points of measurement and across different tasks performed by the same person. The results indicate that EVS/EKS is indeed a sensitive measure useful for process-oriented research. Finally, some methodological challenges and procedures for both interpreting and translation are discussed.

Keywords: time lag, ear-voice span, eye-key span, cognitive processing, process research

1. Introduction

Time lag in interpreting and translation is defined as the temporal delay between source text (ST) input and target text (TT) output. It has been suggested as a variable reflecting temporal characteristics of processing (Pöchhacker 2004: 117) and cognitive load (Treisman 1965; Shlesinger 1998). In translation process research, time lag, defined as the temporal delay between reading source text and producing target text,

is a new measure on which very little research has been done. In simultaneous interpreting (SI), on the other hand, there is a long tradition for measuring time lags between the speaker's input and interpreter's output. The first SI study was carried out in the 1950s (Paneth 1957). Time lag was considered a characteristic of the simultaneity of interpreting, providing an insight into the temporal development of the target speech in relation to the source speech. There seems to be an agreement that the average time lag in simultaneous interpreting is roughly between 2–5 seconds, extending up to around 10 seconds, although at the cost of considerable decline in accuracy (Lee 2002). These findings seem to generally hold (Pöchhacker 2004) across different types of interpreted texts, different interpreters and different languages (including sign languages). So far, researchers measured time lag to establish the basic temporal range, and possibly an upper and lower limit on interpreters' capacity to process the input. Such descriptive studies are important, but it is equally important to probe further. The large variation in time lag suggests that time lag reflects cognitive processing, but little research has been carried out so far to specify the processes reflected in time lag. A number of studies have considered various external factors determining time lag, such as speaker delivery rate (Barik 1973; Lee 2002) or text type (Lamberger-Felber 2001; Podhajská 2008). Time lag is also assumed to be determined by internal (processing) factors, such as strategies (Gile 1997) or segmentation of the input (Goldman-Eisler 1972), but little evidence has been provided so far.

The present paper proposes that time lag has a potential as a temporal variable for process-oriented research in both interpreting and translation. Research into cognitive processes in general relies on two basic variables: speed and accuracy. Process research in translation and interpreting has for long used accuracy as a dependent variable, but a measure of speed of processing has been missing. We wish to investigate the possibility that time lag offers the necessary properties to serve as such a temporal variable. Interpreting research has demonstrated that time lag varies when measured on interpreted output produced on the basis of e.g. different input delivery rates (Gerver 1969/2002), different tasks (such as interpreting with and without text) and also among individual interpreters (Lamberger-Felber 2001), and a similar variation has recently been found in translation (Dragsted 2010). These findings suggest that time lag is indeed a sensitive variable reflecting interpreters' and translators' cognitive processing. Crucially, time lag is an observable variable and a fairly objective one. Process-oriented research is in sore need of such variables: objective, interpretable variables allowing comparison across studies and their follow-up. Before a variable can be used for research in practice, a clear methodology of its measurement needs to be developed. This is the goal of the present paper.

In what follows, we will review available literature on time lag measurement, focusing on the methodology and subsequently present our own analyses of time lag data and discuss issues such as measuring equipment, measuring procedure, points of measurement and quantitative analyses. The aim is to help other researchers make decisions about the use of time lag as a dependent variable in process-oriented research, raise awareness of some methodological difficulties and inherent hurdles, contribute

to improved comparability of results across studies, and eventually develop time lag into a metric applicable in a wide range of research designs aiming to study cognitive processes in interpreting and translation.

2. Time lag measurement in previous interpreting and translation studies

In translation, time lag has traditionally been explored by means of pause detection and analyses. Pauses have been used as the primary indicator of cognitive processing, interpreted as indicators of hesitation and boundaries between text production units or, in translation, translation units or segments (Butterworth 1980; Schilperoord 1996: 11; Jakobsen 1998; Dragsted 2005). More recently, with the integration of eye-tracking technology in Translation Studies, the concept of ear-voice span known from interpreting has inspired a similar measure in translation process research: the eye-key span (EKS) defined (Dragsted and Hansen 2008; Dragsted 2010) as the time which elapses between visual fixation of the ST input and the typed production of the TT output. Since the eye-key span is a very recent variable with little previous research, this section will focus on methodological aspects of time lag measurement in interpreting.

In interpreting, time lag, also known as ear-voice span (EVS) or *décalage*, was first measured in the 1950s (Paneth 1957/2002), and a number of other studies reported analyses measuring time lag. Our aim in this section is to review existing studies and compare methods used for time lag measurement. The studies reviewed include Treisman (1965), Oléron and Nanpon (1965), Gerver (1969), Goldman-Eisler (1972), Barik (1973), Anderson (1994), Lamberger-Felber (2001), Lee (2002), Kim (2005) and Podhajská (2008). The comparison of methods is made primarily in terms of unit of measurement and point of measurement and the possibility to replicate the method, i.e. whether the study contains sufficiently detailed description of method to allow other researchers to use the same method. A summary is provided in Table 1.

The initial trend of measuring time lag as the number of words the interpreter lags behind the speaker was gradually replaced by time units, which also allows better comparability of findings across studies. There is considerably less agreement on the point of measurement. Table 1 clearly shows the wide range of choices of where to measure time lag. This is understandable from a methodological point of view: matching corresponding segments of the source text and the target text is not a trivial matter in interpreting. The original speech often undergoes drastic changes in the interpreted output such as linguistic restructuring due to target text language constraints (e.g. grammatical and syntactical structures), strategic decisions (e.g. substituting several specific elements by a more general one; see e.g. Mackintosh (1985) on the application of macro strategies), or simple errors, omissions and additions. To overcome this inherent difficulty, researchers must choose a method for matching the source text and the target text. For example, Oléron and Nanpon (1965) selected only

Table 1. Methodological overview of previous time lag studies

Author	Year	Unit	Point of measurement	Replicable?*
Treisman	1965	words converted to seconds	every 5 seconds	yes
Oléron and Nanpon	1965	seconds	words with literal translation	yes
Gerver	1969	words	every 5th word of source text	yes
Goldman-Eisler	1972	syntactical units	unclear	no – insufficient detail of method
Barik	1973	seconds	every 5 seconds	yes
Anderson	1994	seconds		no**
Lamberger-Felber	2001	seconds	beginning of segments where at least one interpreter omitted more than 15 words	yes**
Lee	2002	seconds	beginning of sentence	yes
Kim	2005	seconds	beginning of sentence	yes
Podhajská	2008	seconds	unit of meaning (after Lederer 2002)	yes

* This column refers to whether the available study contains sufficiently detailed method section to allow replication.
** Reference to the author's (unpublished, possibly not available in English) thesis with further details of procedure.

words with literal translation output equivalents while Barik (1973: 246) made a more ecologically valid decision to consider also those segments where the "correspondence (…) was based on content rather than simple lexical equivalence". Other researchers chose points of measurement specific to the particular study, such as segments which turned out to be problematic for interpreters (Lamberger-Felber 2001) or segments divided into units of meaning (Podhajská 2008). Methodologically, the crucial question is whether it matters, i.e. whether the different methods of time lag measurement yield different results. We have explored this question in the following analysis.

2.1 Analysis 1: Comparison of previously used methods of time lag measurement in SI

A sample of methods used in previous research was applied to compare the resulting time lag distributions. The methods used by Treisman (1965), Barik (1973) and Lee (2002) were used, as the studies offer sufficient detail to allow replication on a small data set.

2.1.1 *Method*

The materials consisted of an 18-minute speech in English that was interpreted by a professional interpreter and native speaker of Czech with 6 years of experience under laboratory conditions. The speech was on the topic of corporate behaviour in relation to human rights. The speech was read by a native English speaker and interpreted simultaneously into Czech. A 5-minute segment from the middle of the speech was selected for analysis, and the content of the segment was transcribed into two separate files (ST and TT); other elements (voiced hesitation, prolonged pronunciation, intonation, filled pauses) were not transcribed. The speech and its interpretation were recorded on a two-track digital recorder and analysed using digital audio editing software Adobe Audition 1.0[1] (see the final section on methodological observations for a more detailed description of the software).

The three methods of time lag measurement were applied to the interpreted speech. Treisman and Barik measured time lag every five seconds, while Lee measured it at the beginning of each sentence. Treisman measured time lag in the number of words by which the interpreter lagged behind the speaker, Barik and Lee used seconds as the unit of measurement.

(a) *Treisman (1965)*

A cue was inserted into the two-track recording each five seconds, and the point at which the cue was imposed was marked in both the ST and the TT transcriptions. The word on which the cue fell in the TT was traced back to the ST transcription and marked. The correspondence was based on meaning rather than on exact literal translation. For example, when the cue marked the interpreted word *základem* (basis), the corresponding source text expression read *in the heart of*. The word *heart* was thus marked, disregarding the other grammatical words. This approach was applied frequently, as Czech tends to include grammatical markers (gender, tense, cases) in the content word. The distance in the number of words was then calculated between the cue mark (the timestamp) in the source text transcription and the word corresponding to the interpretation. In the example in Figure 1, the 5-s cue is marked in both the ST and the TT (vertical bar). In the TT, the word cued was "v" (preposition "in"). The

ST	**In** some cases, criminal or civil liability may be clear.	For instance on complicity in war crimes …
TT	… zemí však k tomuto nepřihlíží. *[… countries, however, do not take this into consideration.]*	**V** některém případě je jasné, že existuje občansko či trestněprávní odpovědnost. *[**In** some case it is clear that there exists civil or criminal liability.]*

Figure 1. Example of time lag measurement following Treisman (1965)

1. Adobe Systems Incorporated. 2003. Adobe Audition [Version 1.0]. http://www.adobe.com/.

word was traced back to the ST (both are marked in bold). The distance between the 5-s cue (the vertical bar) in the ST (the moment the interpreter pronounced the word "in") and the corresponding ST word ("in") was counted – in this example, it is ten words. Following Treisman, the distance in words was then converted into distance in time on the basis of the input delivery rate. In the example, the delivery rate was 116 words per minute, so the distance of ten words would be equivalent to a time lag of 5.2 s. If no correspondence between the word marked in the target text and the source text could be established, this resulted in a missing value.

(b) *Barik (1973)*
The recording with inserted cues and marked transcriptions generated for the Treisman analysis were used for the Barik analysis. The source text words corresponding to the marked interpretation were then identified in the recording, a second cue was inserted at these points, and the distance between the two cues measured in terms of time. If no correspondence between the word marked in interpretation and the source text could be established, this resulted in a missing value.

(c) *Lee (2002)*
In the two-track recording, beginnings of sentences in the source text were identified and marked with a cue. A second cue was inserted at the beginning of the corresponding sentence in the interpretation. The sentences were treated as units, and a cue marking a sentence in the interpretation was inserted even if only a part of the sentence was interpreted. Thus only sentences omitted in their entirety were registered as missing values.

2.1.2 *Results and discussion*

The mean, median and range values are presented in Table 2. The differences between the mean and median values among the different methods were not statistically significant. This suggests that the various methods produce comparable results and that the choice of method – the unit or point of measurement – does not substantially alter the results. It is, however, notable that the minimum values differ by up to 2 s and the maximum values by up to 1.5 s. The mean/median values are comparable, but specific research into individual items may yield different results depending on the method used. It is therefore worthwhile to examine methodological properties of time lag measurement in more detail and the potential of time lag in process research.

Table 2. Descriptive statistics for time lag measurement using the Treisman, Barik and Lee methods

Author	Items	Valid items	Missing values	Mean time lag (seconds)	Median time lag (seconds)	Range (seconds)
Treisman	59	49	10	4.66	4.65	0.0–9.6
Barik	59	49	10	4.78	4.37	0.2–11.1
Lee	38	31	7	4.60	4.74	2.2–9.5

3. Methodological explorations of time lag measurement

Time lag in translation and interpreting has so far been used as a temporal measure of the process, but has been little explored as a dependent variable reflecting cognitive processing. In order to investigate whether time lag can serve as a reliable dependent variable in translation and interpreting process research, time lag patterns in various contexts need to be considered in more detail. In the following sections, we present our own analyses of time lag in translation and interpreting under varying conditions. The analyses will be presented in two sets, one for translation and one for interpreting, as each mode has a number of specific methodological issues. The analyses for each set address four questions: (1) Does time lag vary over the course of an individual's performance or is it stable throughout the task? (2) Does time lag vary among different individuals or are there shared patterns? (3) Does time lag vary at specific points in a single task? (4) Does time lag vary between different tasks? The rationale for exploring these four questions is to gain a deeper understanding of the cognitive effort of transforming SL input into TL output, and how such effort appears to vary in the course of a given task, across different tasks and across different individuals. In a wider perspective, comparing and contrasting time lags in translation and interpreting contributes useful insights into the coordination and transformation process, which is central to understanding translation and interpreting processes.

4. Time lag measurements in translation

4.1 General method of time lag measurement in translation

Common for all translation experiments reported in this section is that experimental process data were collected using an integrated set of eye-tracking and keystroke logging technologies. Eye-tracking was performed using a Tobii 1750 remote, free-head movement eye tracker, which runs at a sampling-rate of 50 Hz and uses binocular eye tracking.[2] The texts were presented on a TFT display with a resolution of 1280 × 1024 pixels. Participants sat 60 cm from the screen, and were calibrated with a standard 9-point grid. Keystrokes were logged with Translog (Jacobsen and Schou 1999), which presents the full source text from the outset in the top window of the screen and allows the translator to write the target text in the bottom window, and which registers all keystrokes, mouse movements and pauses.

Data from the eye-tracker were analysed with ClearView, a program for gaze analysis. Additionally, fixations in the source text window of the screen were automatically mapped to words with the gaze-to-word mapping (GWM) tool developed in

2. Tobii 1750, see http://www.tobii.se.

the course of the Eye-to-IT project.[3] A study by Dragsted and Hansen (2008) reports that the mapping performed by GWM is subject to considerable uncertainty, with detection rates (fixations mapped correctly to words) at about 80%. Jensen (2008) reports detection rates of between 65% and 75%. In order to be able to efficiently and systematically analyse the eye-key span (EKS), an eye-key manager was developed as an add-on script[4] to our integrated set of eye-tracking and keystroke logging technologies, allowing us to couple ST fixations with equivalent TT production units and retrieve EKS values.

Due to the instability of the gaze-to-word mapping, the eye-key span values described below were calculated on the basis of a selection of thirty word pairs (ST and TT) from each participant for which we had reliable data from GWM. The thirty data points were not the same for all participants, since our main concern in selecting them was reliable mapping from gaze to word, which varied greatly across the participants. This fairly arbitrary selection criterion resulted in a wide distribution across parts of speech with strong predominance of lexical words. Since, according to Clifton, Staub and Rayner, "how long readers look at a word is influenced by the ease or difficulty associated with accessing the meaning of the word" (Clifton Staub and Rayner 2007: 248), one might expect more robust data from difficult words associated with longer fixations, also in translation (Sharmin, Špakov, Räihä and Jakobsen 2008: 43). However, we analysed eight randomly selected words for which we had reliable gaze-to-word mappings from four participants in each group and found that at least three of the words did not appear problematic, neither from a subjective point of view (does this word seem difficult to translate), nor from a more objective point of view (does it attract longer-than-normal fixations, more refixations or longer eye-key spans). Thus the thirty selected word pairs do not appear to be particularly difficult.

For each word pair we deducted the timestamp in ms of the onset of the fixation on a ST word from the onset of the production of the equivalent TT word. For instance, if the timestamp of a fixation on the ST word *better* was 10,000 ms, and the timestamp of the production of the first letter in the equivalent *bedre* was 10,800 ms, the eye-key span would be 800 ms. ST words (in particular) were often associated with several refixations; hence, EKS values can be computed either as the time lag from the first time the word was fixated in the ST (EKS from first fixation) or from the last fixation on the ST word before producing the TT word (EKS from last fixation), or in principle from any fixation in between, see below.

The EKS analyses presented below draw on data from several translation process studies of translation students and/or professional translators carried out at the

3. See http://cogs.nbu.bg/eye-to-it for a description of the project and http://www.cs.uta.fi/~oleg/gwm.html for a description of the gaze-to-word mapping tool developed by Oleg Spakov.

4. We are grateful to Laura Winther Balling (Copenhagen Business School) for helping us write the script.

CRITT[5] centre at Copenhagen Business School (CBS), using the methodology described in this section. We will give examples of experiments with and analyses of EKS in written translation and raise a number of methodological issues. At the end, we will discuss some of these challenges and point to strengths and weaknesses in our design and methodology and make suggestions as to how to counteract methodological flaws.

4.2 Analysis 2: Individual EKS variability (intra-subject analyses)

Using data from four students and four professional translators translating the same text, we measured individual eye-key span patterns and observed how these may vary during the process of translating a 100-word text. The analyses showed (1) that the degree of intra-subject variability differed across the two groups of subjects studied (students vs. professionals), and (2) that variability was higher for EKS from first fixation than for EKS from last fixation.

In the examples in Figure 2, the professional translator's EKS pattern generally remained more stable throughout the translation task than that of the student with respect to EKS from the first as well as from the last fixation.[6] Notably, the EKS from the last fixation remained consistently below five seconds for the professional, whereas the students had "fluctuations" of up to 40 seconds, see Figure 3.

The 40-second peak around word 10 may be the result of missing data rather than an indication that the subject stored the source text word in working memory for 40 seconds, which seems unlikely due to working memory constraints. In other words, there is reason to believe that there was a refixation on the word prior to the production of the equivalent, but that this fixation has not been registered by the eye-tracker or has not been mapped correctly by the GWM tool, since the tools are known to exhibit some imprecision (Jensen 2008). Generally, one should pay attention to outliers in the EKS from the last fixation and consider whether they should be excluded from the mean values. Due to working memory constraints, one would not expect large variation on this parameter (for differences between the two groups, see below).

As for the EKS from the first fixation, however, one can expect significant variation. In the current study, translators in both groups sometimes fixated ST words long (in some cases 1.5 minutes) before they started producing target language equivalents.

---

5. Center for Research and Innovation in Translation and Translation Technology at CBS.

6. The examples illustrate prevalent patterns in the two groups of subjects. In the group of professionals, the EKS from the last fixation, i.e. just before commencing the production of the TL equivalent, ranged from 5 ms to 12.5 seconds (SD between 866 ms and 3 seconds). In the group of students, EKS from the last fixation ranged from 438 ms to 51.5 seconds (SD between 1.5 second and 12.5 seconds). EKS from first fixation ranged from 696 ms to 306 seconds (SD between 54 and 72 seconds) in the group of students, whereas professionals' EKS ranged from 5 ms to 201 seconds (SD between 23 and 41 seconds).

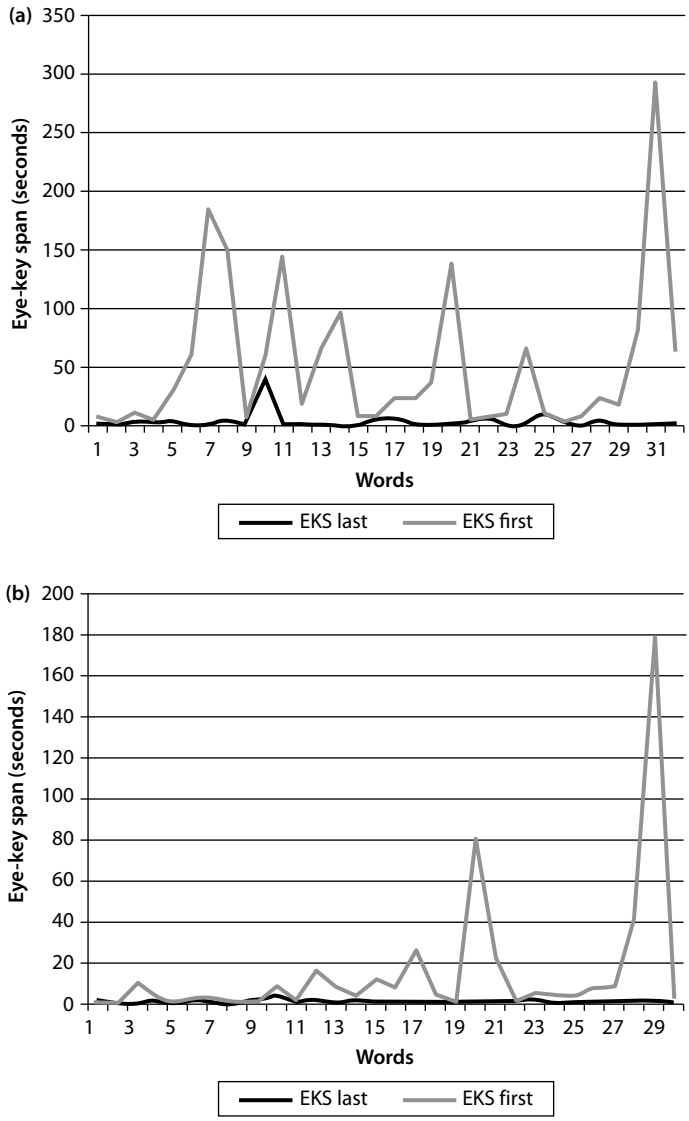


Figure 2. Individual eye-key span patterns in (a) a student, (b) a professional translator

A previous study suggests that problem words such as metaphorical expressions attract attention and are fixated long before they are translated (Dragsted and Hansen 2008: 12). Conversely, the first fixation on a word may occur immediately before production, the EKS from the first fixation coinciding with the EKS from the last fixation. This leads to high intra-subject variation in both groups, but again, student data generally exhibited more variation than data from professionals. Further investigations of EKS should involve analyses of the nature of source text words attracting long/short

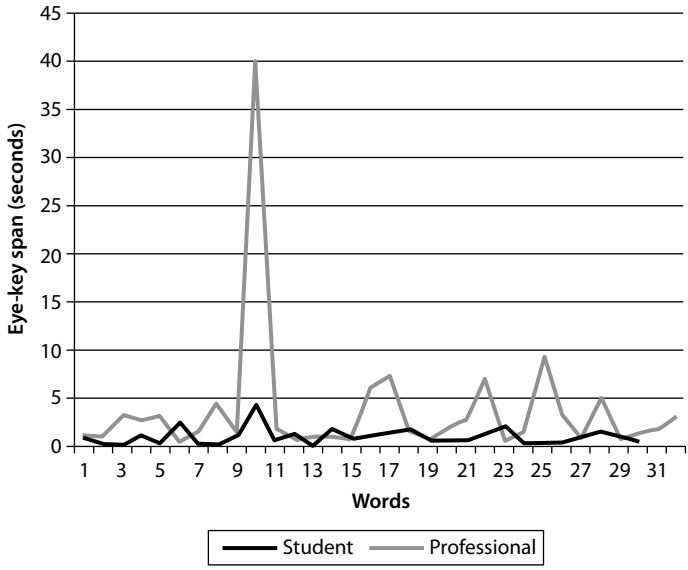


Figure 3. Eye-key span from last fixation

EKS based on syntactic as well as semantic annotation of source and target text. Work is currently being done (at CBS) to arrive at more elaborate representations of eye-tracking and keylogging data.

4.3 Analysis 3: EKS patterns across individuals (inter-subject analyses)

Analyses of EKS across individuals were performed mainly as a comparison between a group of eight M.A. translation students enrolled in the translation and interpreting programme at the Copenhagen Business School and eight professional translators who have studied and practised translation for at least ten years and may be considered translation experts (Ericsson 2000/2001: 213). All participants had Danish as their mother tongue and English as their first foreign language.

Before commencing the actual translation task, participants were allowed time to read the ST. The advantage of isolating the pre-translation phase is that EKS values from the first fixation may otherwise reflect major fluctuations, depending on whether the translator applies this strategy or not, and would distort comparisons between translators who read through a text before translating it and translators who do not.

Table 3 shows the EKS from first and last fixation in professional translators and students. The students' higher values of EKS from last fixation of more than 7 s on average indicate that students spent more time and effort constructing meaning based on the ST and switching modes in order to commence TT production. This also includes the physical effort of finding the right key. In this and earlier studies (Dragsted and Hansen 2008; Dragsted 2010), we have observed major differences in typing skills

Table 3. Descriptive statistics for eye-key span measurement in professional translators and translation students

	Professionals	**Students**
Mean/median EKS from last fixation (seconds)	2.8/2.8 (SD 1.1)	7.2/6.4 (SD 3.9)
Mean/median EKS from first fixation (seconds)	19.7/18.6 (SD 9.9)	58.6/52.0 (SD 12.2)
Fixations per ST word (mean)	3	6

between groups and within groups. Typing skills may have an impact on the EKS and should therefore be taken into consideration. In the section "Comparing EKS in translation and copying tasks" below, we report a first attempt at including this parameter in our analyses by measuring the time span between a ST word fixation and the retyping of the fixated word (in the same language). This should give some indication of typing speed and the relation between typing skills and EKS.

In some cases, the EKS comprised pauses immediately preceding production units, and students tended to have several fixations on words other than the ST unit to be translated between the last fixation on the ST word and the production of its equivalent. A closer analysis of this phenomenon would be interesting as it might provide more insight into short-term memory capacity. Also it might indicate subjects' anticipation procedures and patterns, i.e. to what extent translators read ahead of the word under production and possible differences between translators' and interpreters' working patterns.

The EKS values from the first fixation do not include fixations related to reading through the text before commencing the actual translation tasks since this type of reading activity had been isolated in the pre-translation task as mentioned above. Hence, the high EKS values of up to several minutes from the first fixation particularly in the group of students, seem to be another indication of their exertions to arrive at an understanding of each ST element in order to be able to produce TT equivalents. The long time lags between fixating a ST word for the first time and producing an equivalent was often associated with several refixations (six on average in the whole text, see Table 3) in the group of students, whereas in the group of professionals, a fixation on a ST word frequently resulted in a production unit.

The higher EKS values among students may be taken as an indication of the time it takes for the student to construct meaning based on the ST and to switch mode in order to commence production. It has previously been suggested (Dragsted 2010) that the coordination of reading and writing processes is more *sequential* in students than in professionals: the students read long sequences of the ST before they switch mode and start producing (a sequence of) TT (cf. consecutive interpreters who listen to a sequence of speech and interpret it afterwards). Hence, time lags between reading ST and producing TT are relatively long. The professionals, on the other hand, continuously shift back and forth between ST and TT; coordination/transformation is instant, and hence time lags between reading ST and producing TT are relatively short. The term *integrated* coordination has been used (Dragsted 2010) to cover this

type of behaviour where TT production takes place almost simultaneously with ST reading (cf. simultaneous interpreters who speak as they listen).

4.4 Analysis 4: Eye-key span relative to parts of speech and sentence position

The scarcity of reliable data from GWM and the methodological limitation of having to base the measurement of EKS values on high GWM quality word pairs prevented us from performing extensive analyses of the relationship between specific parts of speech and EKS values. Using analyses of selected EKS values, it was not possible to establish a correlation between for instance longer-than-average or shorter-than-average EKS (either from the first or the last fixation) and specific parts of speech. EKS with durations above or below the average appeared to be dispersed more or less randomly across different parts of speech (verbs, nouns, adjectives and adverbs)[7] and different parts of the text. As already mentioned, work is currently progressing to improve data quality and enable syntactic and semantic annotation of product and process data (Carl and Jacobsen 2009).

There was a tendency in both groups for EKS from the first fixation to become longer towards the end of the task, but this is hardly surprising since the probability that a word has been fixated long before it is produced increases as task time accumulates. There was also a slight tendency, particularly in the group of students, for some words to attract long EKS from the first fixation across all subjects in the group. For instance the adjective *economic* (mean EKS 18.3 sec.) and the verb *offered* (mean EKS 15.7 sec.) were associated with longer-than-average EKS from the last fixation. One might speculate whether these words would have caused problems for the translators. However, since the limited amount of data would only allow us to make speculative conclusions, we have not performed further analyses on these cases.

4.5 Analysis 5: Comparing EKS in translation and copying tasks

In an attempt to single out the effort of transforming source language input to target language output, deducting the effort related exclusively to typing, we have compared the EKS in a translation task with the eye-key span in a copying task. Six professional translators were asked to copy a text in English, which was their first foreign language, and to translate a similar text from English into Danish.

As illustrated in Figure 4, the curves are not parallel, but still indicate that there may be a correlation between EKS in the two tasks. The distance between the two curves may be interpreted as an indication of the translation effort. As mentioned

7. The words were *understandably, reluctant, compromise, threaten, economic, development, incentives, emissions and vulnerable.*

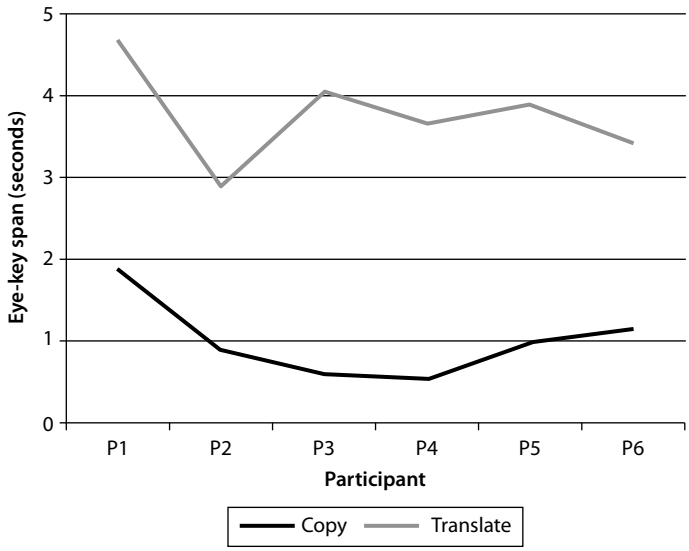


Figure 4. Individual eye-key spans for copying task and translation task

above, differences in typing skills influence eye-key spans. In order to compensate for these differences, one might include copying tasks in EKS experiments or alternatively select subjects on the basis of typing skills, with preference for touch typists. This would have the desired side effect that touch typists generate higher quality gaze data because their eyes are constantly fixated on the screen. However, when selecting student participants in particular, this criterion may strongly reduce the number of eligible subjects.

4.6 Methodological observations

EKS measurement in translation process studies using eye-tracking and keystroke logging technologies is still in its infancy, and results drawing on data from several translation process studies carried out at CRITT indicate that the tools and methods applied need to be tested further and fine-tuned to improve data quality and reliability. Thus the instability of GWM mapping hampers the selection of data points for analysis where the aim is, for instance, a detailed analysis of the relationship between EKS values and specific parts of speech. In the same vein, automatic detection and retrieval of matching ST and TT units of analysis is a major challenge, and the matching of ST and TT words described in the study above was manual and based mainly on lexical units. EKS measurements based on syntactic as well as semantic ST and TT annotation require further research and development, i.e. more elaborate representations of eye-tracking and keylogging data. In a future scenario, integration of features of machine translation and text alignment tools may be explored as candidates for ST/TT unit detection and "matchmaking".

The analyses presented in this paper are based on EKS measurements of first and last fixations, giving an indication of time lags between fixations on ST words and TT equivalents. In an attempt to define/determine the data point of first fixation and generate more reliable and comparable data, a pre-translation reading task for all subjects was included in our experimental procedure. This procedure seems to have improved data comparability between subjects.

However, depending on the research question, it is important not only to distinguish between first and last fixations, but to probe into and break down time lags in order to investigate which words/structures trigger high EKS values. Interpretation of fixations patterns involving all fixations/refixations on data points between first and last fixations is as yet uncharted territory and raises a number of questions about the coordination of comprehension and production processes, i.e. "what goes on" in between the first and the last fixation on a data point before the onset of output production. A further and more detailed breakdown of EKS values may bring out insights into subjects' anticipation procedures and patterns as well as give an indication of their short-term memory capacity.

As stated above, major differences in typing skills, observed between groups and within groups, may have an impact on the EKS values and should be considered, for instance by including a typing task in the experimental design to take into account general time consumption differences, e.g. between professional translators and novices.

5. Time lag measurements in interpreting

5.1 General method of time lag measurement in interpreting

The analyses of simultaneous interpreting presented below addressed the same questions as the translation analyses described above, namely variations in time lag between individual interpreters, time lag variations within individuals, time lag variations at specific points of measurement and time lag variations between different tasks. For all analyses, unless stated otherwise, the following method was applied.

The materials are the same as those used for the initial analysis of previously used method of measurement. One 18-minute speech in English was interpreted by 16 interpreters under laboratory conditions. The speech was on the topic of corporate behaviour in relation to human rights. It was read by a native English speaker and interpreted simultaneously into the interpreters' native language. The source text and its interpretation were recorded on a two-track digital recorder and analysed using digital audio editing software, Adobe Audition 1.0.

In the two-track recording, the points of measurement were identified in the source text and marked with a cue. A second cue was inserted at the beginning of the corresponding segment in the interpretation. If sentence beginnings were selected as points of measurement, sentences were treated as units, and a cue marking a sentence in the interpretation was inserted even if only a part of the sentence was interpreted.

Thus only sentences omitted in their entirety were registered as missing values. The distance between the two cues was measured and time lag calculated in seconds with a 0.1-s precision.

5.2 Analyses 6 and 7: Individual ear-voice span variation (intra-subject analysis) and ear-voice span patterns across individuals (inter-subject analysis)

Several studies have examined extensively the relationship between time lag and other variables, and other authors have reported assumptions and observations from practice. Gile (1997) emphasises the strategic nature of time lag management. Lee (2002, 2003) reports a study in which time lag was related to accuracy and speaker variables. As for accuracy, Lee found that time lags exceeding a certain limit (approximately four seconds) result in a decrease in accuracy. Making a fairly safe assumption that interpreters strive for accuracy, time lag leading to incorrect interpretation cannot be considered a strategic decision by the interpreter. Lee also found a relationship between time lag and speaker variables, such as presentation rate and the number of pauses made by the speaker. That would suggest that time lag is – at least to some extent – driven by external factors (Barik 1973; Lee 2002). If that were the case, multiple interpretations of the same speech should demonstrate similar time lag patterns. This question was explored in the following analysis.

5.2.1 *Method*

The participants were 16 professional interpreters, native speakers of Czech and Dutch (10 and 6 interpreters respectively), with between one and 21 years of experience. In the materials described above, 30 measurement points, all beginnings of sentences, were selected for time lag measurement.

5.2.2 *Results and discussion*

The analysis revealed no common pattern among the interpreters; in fact, individual charts demonstrate high variability both in terms of inter-subject and intra-subject comparisons. Figure 5 shows an example of two interpretations of the same speech in a text with 30 measurement points at the beginnings of sentences. The chart clearly demonstrates that different interpreters have very different time lag patterns. This excludes external factors, such as the speaker's delivery rate, as the sole determinant of time lag. Similarly, interpreters do not keep a constant time lag, suggesting that time lag is not determined solely by the interpreters' constant "preferred" distance from the speaker, but by other factors as well. The lack of similarity between interpreters suggests that the causes of specific time lags are probably a combination of external and internal determinants, such as subjective perception of speech difficulty, language difficulty, delivery rate, familiarity with the topic, accent, etc.

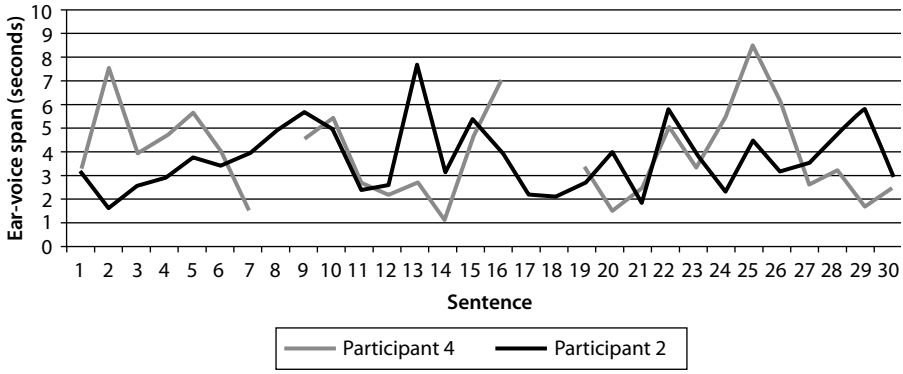


Figure 5. Individual ear-voice spans in two participants

5.3 Analysis 8: Ear-voice span relative to parts of speech

In this analysis, we examined the behaviour of the time lag metric in a more specific context. While the first analysis compared the measurement of average time lags over the whole duration of an interpreted text, the second analysis had a more narrow focus on the measurement of time lag at pre-selected points and time lag differences according to the position in the text of the measurement point. Specifically, we measured time lag at three different points: sentence beginnings, verbs and figures. The choice of the three categories is arbitrary. Sentence beginnings are convenient as they are relatively easy to locate in the recordings, and sentences as units are rarely missing values. Verbs are fairly numerous, and given their prominent syntactical role it is assumed they will be usually interpreted. Figures were selected because they are known to be especially difficult as they cannot be inferred from the context.

5.3.1 *Method*

A single interpretation by a professional interpreter and native speaker of Czech with 10 years of experience was selected from the materials described above. Thirty sentences of three different types were analysed: syntactically complex sentences, sentences containing negative verbs and sentences containing two or three figures. Sentences within each category were matched in terms of length and structure.

5.3.2 *Results and discussion*

Table 4 shows the mean and median values of time lag measured at the various measuring points. A one-way analysis of variance has shown that there are significant differences in time lags in that the time lag measured on figures is significantly shorter than time lags measured on verbs or at the beginnings of sentences. There was no difference between verbs and beginnings of sentences.

Table 4. Descriptive statistics for ear-voice span measurement on different parts of speech

Point of measurement	Items	Valid items	Missing values	Mean (seconds)	Median (seconds)
Sentence	30	30	0	3.6	3.2
Verb	32	24	8	3.8	4.1
Figures	17	10	7	2.6	2.5

The results suggest that it makes a difference at which point in the text time lag is measured, which runs contrary to the outcome of the comparison of the various methods reported in the introductory comparison of previous research. One reason for this difference could be that a time lag measurement aiming at establishing the mean value of time lags throughout the whole interpreted text may include widely disparate points of measurement, as is indeed the case in the Barik and Treisman methodology. Barik and Treisman measured time lag every five seconds and did not differentiate between parts of speech. In the example shown in Figure 1, time lag is measured on a preposition. On the other hand, in our analysis we distinguished between specific points of measurement on the basis of their linguistic properties, which may require different types of processing. The present analysis seems to support the notion that time lag can be used as a sensitive metric, distinguishing between the various sub-processes involved in simultaneous interpreting. Both approaches are valuable. The overall mean time lag measurement could be more useful in cases where different interpreting situations are compared on a more global level, such as the comparison of interpreting of prepared and impromptu speeches (Podhajská 2008). The more sensitive approach illustrated here may be of potential use in experimental studies, focusing on different sub-processes at a more local level, such as comparison of specific structures or difficult terminology. These suggestions are only tentative at this point; more research will be needed to assess the sensitivity of time lag as a useful metric in this type of process research.

5.4 Analysis 9: Comparison of ear-voice span in interpreting and shadowing

Shadowing is a task in which participants are asked to simultaneously repeat verbatim an auditorily presented text, either in a native or a foreign language. This task has been used as part of simultaneous interpreting aptitude tests (e.g. Moser-Mercer 1985; Lambert 1991), as it requires the division of attention between listening and speaking. It is interesting to explore whether there is any relationship between shadowing and interpreting, two tasks which both require simultaneous processing, albeit of a different nature.

5.4.1 *Method*

The data consist of a serendipitous corpus of multiple task performance by student interpreters. Twelve students shadowed an authentic conference speech in English and interpreted an artificially created speech from English into Czech, the students' native language. The data were collected as part of an aptitude exercise, administered to translation and interpreting students at the end of a one-year consecutive interpreting course. It needs to be pointed out that this was the students' first time in the interpreting booth – both shadowing and simultaneous interpreting were novel tasks in which they had no practice. The two speeches were selected to serve the purposes of the aptitude exercise, and no attempt was made to make them comparable. The shadowing speech was on the topic of computer crime, spontaneously delivered by a native English speaker. The speech was 4 min long. The interpreting speech was a short narrative text about two friends spending a Sunday in the park. The speech was 4 min 45 s long. Four of these participants (9–12) participated in another study seven years later. At that time they all worked as professional interpreters and had one to seven years of professional experience. Recordings from this latter study were included in the present analysis. The data, described above, consisted of a simultaneously interpreted speech on the topic of corporate behaviour in relation to human rights. Time lag was measured at beginnings of sentences for all texts. There was a total of 24 measurement points in the shadowed speech, 20 measurement points in the interpreted speech and 30 measurement points in the speech interpreted by the participants as professional interpreters.

5.4.2 *Results and discussion*

Figure 6 shows the individual median time lag values for each participant and task. The overall median values were 1.39 s in shadowing and 4.03 s in interpreting. This is consistent with previous findings of shadowing time lag being shorter than interpreting time lag (Treisman 1965; Anderson 1994) and the assumption that longer time lag in interpreting reflects more elaborate processing. From an individual differences perspective, however, there seems to be no relation between the time lag on the shadowing task and the interpreting task, i.e. shorter time lag in shadowing does not mean shorter time lag in interpreting. It must be borne in mind, however, that the participants were students attempting interpreting for the first time. As for the comparison between the student and professional interpreter performance of participants 9–12, three out of four demonstrate a similar time lag on both interpreting tasks completed at different points in time (one as students, one as professional interpreters). The data are too limited to allow any assumption as to stability of time lag over time, but raise an interesting question as to whether there is a "personal" time lag. This could be a preferred distance from the speaker related to individual cognitive makeup and processing preferences.

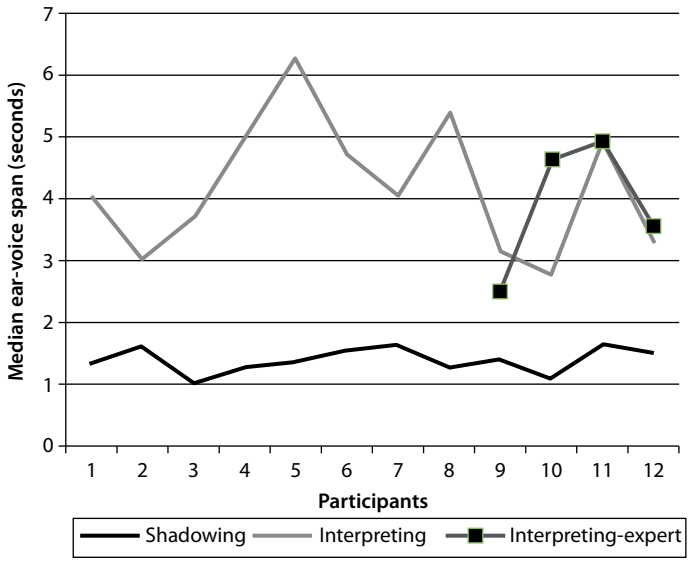


Figure 6. Individual median ear-voice spans in shadowing and interpreting

5.5 Methodological observations

In order to support replicability of time lag studies, in this section we will consider some of the methodological issues we have encountered and tried to resolve.

Equipment. Technology development has dramatically changed time lag measurement methods. The first studies used a stop watch and equipment converting sound mechanically into a graphic representation (polygraph). Nowadays, audio-editing software is available allowing much more precise measurements and analyses. The software used for interpreting analyses in this paper is Adobe Audition 1.0, allowing editing of stereo, i.e. two-track recordings. Each track is loaded into a separate channel, and both channels can be processed at the same time, or only one channel can be enabled. The recording can be "zoomed in" to allow file processing at a more detailed level. Figure 7 shows several screenshots of the same recording with different zoom. While the sound waves can be zoomed in almost infinitely, this is not necessarily practical. In Figure 7 (a) the whole recording is visible in one screen (10 min 40 s), the grid unit is 30 s. Importantly, the cursor moves in steps of 625 ms, which is fairly imprecise. In (b), the recording is zoomed, the screen shows 1 min 27 s of the recording, the grid unit is 5 s, and the cursor moves by 85 ms, in (c) the screen shows 40 s of recording, the grid unit is 2 s and the cursor moves by 36 ms. Finally, in (d), the screen shows 9 s of recording, the grid unit is 0.5 s and the cursor moves by 8 ms.

Points of measurement still need to be detected manually by a human, and the human ear cannot perceive with millisecond precision. Segments lasting around 0.2 s

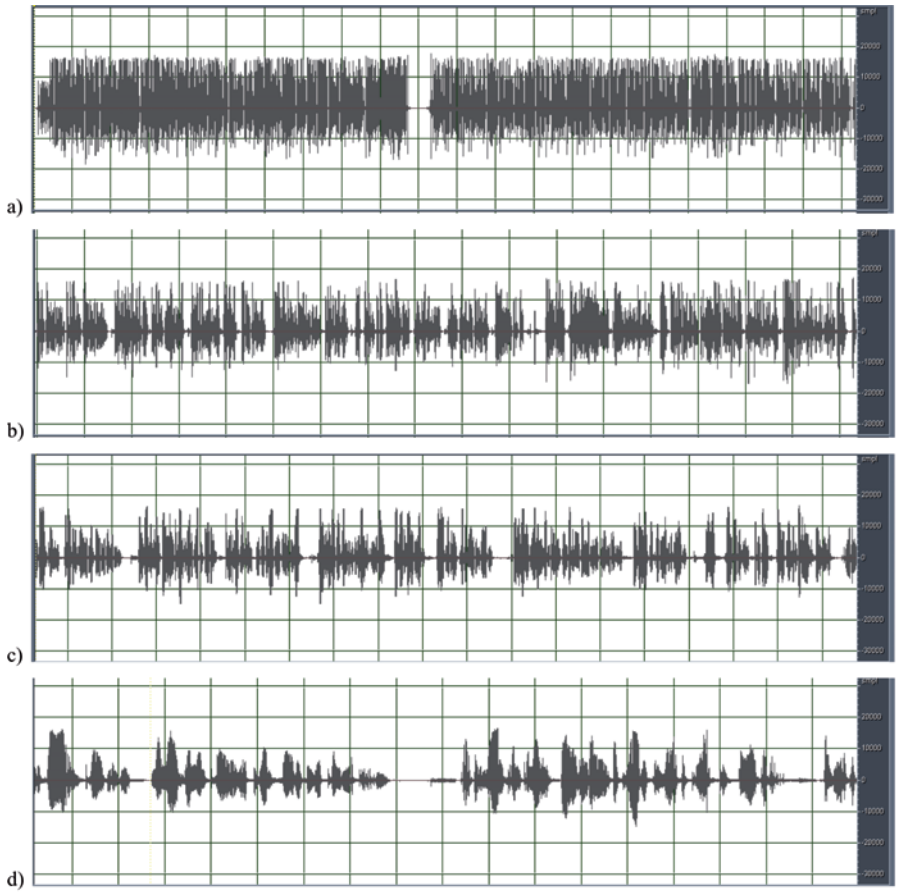


Figure 7. Screenshots of a recording in Adobe Audition with varying zoom: (a) no zoom, grid unit 30 s, (b)–(d) zoomed, grid unit (b) 5 s, (c) 2 s, (d) 0.5 s

contain approximately one syllable and below this level, measurements dependent on human perception tend to be inaccurate. Also, managing a recording file of 20 min, where only 10 s are visible at a time, becomes confusing: such a speech consists of 120 screens (six ten-second screen views per minute, times 20 minutes) and it is not easy to move back and forth and keep track of them. A workable setup adopted for analyses in this paper is a zoom to screen showing approximately 40 s of recording with a grid scale of 2 s (illustrated in (c) in Figure 7). At this zoom level, the cursor can be adjusted with around 40 ms precision (0.04 s). Considering the basic time lag measurement unit of 0.1 s, this seems to be adequately accurate for time lag analysis. Different setups may be required in cases where the researcher is interested in local processes related to a single sentence or word, rather than measurements across a whole speech.

Adobe Audition allows the user to save recordings in a variety of formats. When recordings are saved as .wav files, certain modifications and non-audio information can be included in the file. For time lag measurement purposes, cues can be inserted at the points of measurement and saved in the file without altering the recording in any way. The cues can be preserved for future reference or further work on the file. Furthermore, cues can be inserted as single points or as cue ranges. By positioning the beginning of a cue range at the point of measurement on one channel (original speech) and the end of a cue range at the point of measurement on the other channel (interpreting), the software will automatically calculate the duration of the cue range – saving the researcher the need for a manual calculation of the time lag.

Other researchers have reported use of other software, such as Sony Sound Forge (Kim 2005) or Creative SoundWave (Lee 2002). Both seem to provide a similar level of precision of measurement, with the difference that the original speech and its interpreted output have to be loaded as two separate files, which makes careful synchronisation essential. The issue of synchronisation is of special importance where recordings are made on traditional audio tapes. Podhajská (personal communication) found that due to slight differences in the speed at which tapes are played (during recording or playback), the difference in duration of the same file (such as multiple recordings of the original speech) can differ by several seconds. It seems therefore important that time lag measurements are made separately on each ST/TT pair, rather than making time lag measurements once on a "master" source text recording and applying these to all subsequent TT recordings.

Unit of measurement. Availability of software with millisecond measurement precision is tempting. At this stage, however, the temptation should probably be resisted. There are two main reasons for that. First, as pointed out above, no matter how sensitive the software, analyses still need to be done by humans, and measurement is thus subject to the limitations of human perception. Auditory or visual perception of an onset of the item of interest with one-millisecond accuracy is impossible. Secondly, even if the sound wave clearly shows a boundary between two bursts of speech, these will usually occur where there is a pause. Spoken language consists of a continuous speech signal, words naturally blend together, and detecting an exact boundary between individual words in the sound wave is difficult. Attempting to measure with a one-millisecond precision will most likely lead to inaccuracy. Lee (2002) points out that even though precise software was used, time lag was measured in units of 100 ms (0.1 s). This is the preferred unit in other studies as well, and our own analyses support this approach.

Statistical analyses. Most studies report the mean value of time lag per interpreter and/or speech. Mean is a summary statistics, providing one, "average", value. It is, however, sensitive to extreme values. If most time lag measurements are, for example, around 2.5 s, and only a few values are well above this value, the resulting mean can be distorted. This was already pointed out by Oléron and Nanpon (1965/2002: 47), who opted for a comparison of median values. Podhajská (2008) reported mean values of time lag for two speeches, each interpreted by 12 interpreters. Re-analysis of the raw

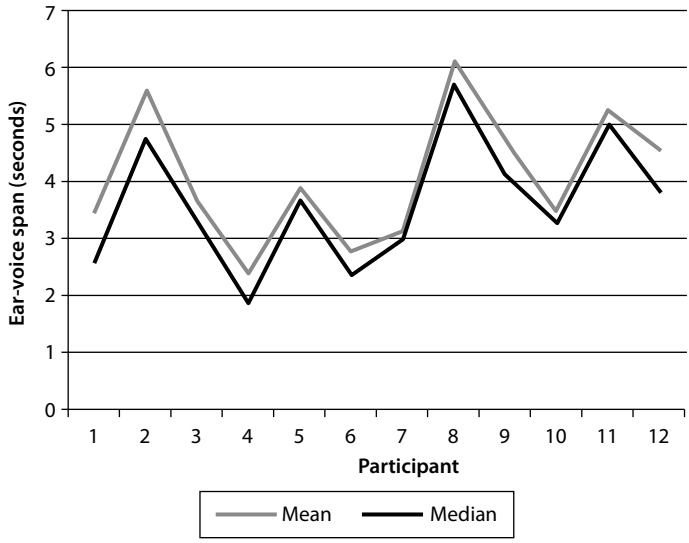


Figure 8. Mean and median ear-voice span – raw data from speech 2, Podhajská (2007)

data available in Podhajská (2007) revealed differences illustrated in Figure 8 and Table 5. Median values were consistently lower than mean values. A paired two-sample t-test shows that the difference is statistically significant ($t(23) = 9.94$, $p < 0.001$). This suggests that the asymmetrical distribution of time lag measurements is not suitable for mean calculation, as the mean provides a distorted picture of the time lag due to its being influenced by outliers and extreme values. It is advisable that researchers report not only the mean values, but also the standard deviation, median and range.

Table 5. The difference between individual and overall mean and median ear-voice span

Participant	Mean (seconds)	Median (seconds)	Mean-median (seconds)
1	3.4	2.6	0.8
2	5.6	4.8	0.8
3	3.7	3.3	0.4
4	2.4	1.9	0.5
5	3.9	3.7	0.2
6	2.8	2.4	0.4
7	3.1	3.0	0.1
8	6.1	5.8	0.3
9	4.7	4.1	0.6
10	3.4	3.3	0.1
11	5.3	5.0	0.3
12	4.5	3.8	0.7
	Mean	Median	Median
Total	4.1	3.5	0.4

6. Conclusion

The analyses presented in this paper aimed at exploring various methodological issues related to time lag in translation and interpreting. Due to the different nature of the two tasks, there are many differences between the two modes. Each mode relies on different technology and equipment and faces different methodological challenges. Interpreting research is still heavily dependent on manual measurement. Translation research makes use of sophisticated eye-tracking and keylogging technologies, but also here, analyses still depend greatly on subjective interpretations on the part of the researcher. The use of EKS as a measure in translation process studies, using eye-tracking and keystroke logging technologies, is still at the experimental stage, and the tools and methods applied need to be tested and refined in order to improve data quality and reliability. Notably, the instability of mapping fixations to words hampered a detailed analysis of the relationship between EKS values and specific parts of speech. Overall, we need to use the measure with extreme caution and as a source of evidence supplementary to product data and other process data.

However, there are many similarities between time lag measurements in translation and interpreting. Both modes struggle, in their own ways, with missing data, which prevents design of carefully manipulated experiments targeting specific phenomena. On the other hand, the present analyses for both translation and interpreting clearly demonstrate that time lag is a variable with a clear promise to allow glimpses into the translator's and interpreter's mind. We found the time lag measure to be highly sensitive to a variety of factors related both to the task and to the translator/interpreter. Findings from both translation and interpreting show that time lag is highly variable within a single performance of one translator/interpreter. In translation experiments, professional translators' EKS values generally remained stable throughout the translation task, whereas the students had more variability during the course of the task with respect to EKS from the first as well as from the last fixation. The fluctuations in the students indicate that the flow of production is more often disrupted by "stumbling blocks" requiring a larger cognitive effort reflected in longer time lags. In interpreting, individual performances by professional interpreters showed large variations in time lag throughout the task, which suggests that the demands of the task are not the same at all times.

Time lags vary also across different individuals in both translation and interpreting. Professional and novice translators exhibit different behaviour when it comes to the EKS from the first as well as the last fixation on a ST word. High EKS among students indicate that they spend more time and effort constructing meaning and switching from comprehension to production mode. Comparison of several interpreters interpreting the same text showed that time lag is not the same for all of them at specific points, which strongly suggests that objective factors, such as speaker variables, are not the sole determinant of time lag, and that other, more subjective factors play a role. Measurements taken at different points, such as on specific parts of speech, showed different time lag, and comparison of translation/interpreting with

more simple tasks demonstrated shorter time lag for the simple tasks, such as text copying and shadowing. All of these are exciting results suggesting that when used with caution and careful method, time lag could serve as a temporal variable reflecting processing speed, providing an opportunity to further probe the underlying cognitive processes.

References

Anderson, L. 1994. "Simultaneous interpretation: Contextual and translation aspects." In *Bridging the Gap: Empirical Research in Simultaneous Interpretation*, S. Lambert and B. Moser-Mercer (eds), 101–120. Amsterdam/Philadelphia: John Benjamins.

Barik, H. C. 1973. "Simultaneous interpretation: Temporal and quantitative data." *Language and Speech* 16 (3): 237–270.

Butterworth, B. 1980. "Evidence from pauses in speech." In *Language Production. Volume 1. Speech and Talk,* B. Butterworth (ed.), 155–176. London: Academic Press.

Carl, M. and Jakobsen, A. L. 2009. "Objectives for a query language for user-activity data." In *Natural Language Processing and Cognitive Science (NLPCS 2009) Conference*, 6–8 May, 2009, Milan, Italy.

Clifton, C., Staub, A. and Rayner, K. 2007. "Eye movements in reading words and sentences." In *Eye Movements: A Window on Mind and Brain*, R. P. G. van Gompel, M. H. Fischer, W. S. Murray and R. L. Hill (eds), 341–371. Amsterdam: Elsevier.

Dragsted, B. 2005. "Segmentation in translation – differences across levels of expertise and difficulty." *Target* 17 (1): 49–70.

———. 2010. "Coordination of reading and writing processes in translation: an eye on uncharted territory." In *Translation and Cognition*, G. Shreve and E. Angelone (eds), 41–62. Amsterdam/Philadelphia: John Benjamins.

Dragsted, B. and Hansen, I. G. 2008. "Comprehension and production in translation: a pilot study on segmentation and the coordination of reading and writing processes." In *Looking at Eyes. Eye-Tracking Studies of Reading and Translation Processing* [*Copenhagen Studies in Language 36*], S. Göpferich, A. L. Jakobsen and I. M. Mees (eds), 9–29. Copenhagen: Samfundslitteratur.

Ericsson, K. A. 2000/2001. "Expertise in interpreting: an expert-performance perspective." *Interpreting* 5 (2): 187–220.

Gerver, D. 1969/2002. "The effects of source language presentation rate on the performance of simultaneous conference interpreters." In *The Interpreting Studies Reader,* F. Pöchhacker and M. Shlesinger (eds), 52–66. London/New York: Routledge.

Gile, D. 1997. "Conference interpreting as a cognitive management problem." In *Cognitive Processes in Translation and Interpretation*, J. H. Danks, G. M. Shreve, S. B. Fountain and M. K. McBeath (eds), 196–214. London/Thousand Oaks/New Delhi: Sage.

Goldman-Eisler, F. 1972. "Segmentation of input in simultaneous translation." *Journal of Psycholinguistic Research* 1 (2): 127–140.

Jakobsen, A. L. 1998. "Logging time delay in translation." In *LSP Texts and the Process of Translation* [*Copenhagen Working Papers 1*], G. Hansen (ed.), 71–101. Copenhagen: Copenhagen Working Papers in LSP.

Jakobsen, A. L. and Schou, L. 1999. "Translog documentation." In *Probing the Process in Translation: Methods and Results* [*Copenhagen Studies in Language 24*], G. Hansen (ed.), 151–186. Copenhagen: Samfundslitteratur.

Jensen, C. 2008. "Assessing eye-tracking accuracy in translation studies." In *Looking at Eyes. Eye-Tracking Studies of Reading and Translation Processing* [*Copenhagen Studies in Language 36*], S. Göpferich, A. L. Jakobsen and I. M. Mees (eds), 157–174. Copenhagen: Samfundslitteratur.

Kim, H. 2005. "Linguistic characteristics and interpretation strategy based on EVS analysis of Korean-Chinese, Korean-Japanese interpretation." *Meta* 50 (4): http://id.erudit.org/iderudit/019846ar (Accessed February 19, 2011).

Lambert, S. 1991. "Aptitude testing for simultaneous interpretation at the University of Ottawa." *Meta* 36 (4): 586–594.

Lamberger-Felber, H. 2001. "Text-oriented research into interpreting. Examples from a case-study." *Hermes, Journal of Linguistics* 26: 39–63.

Lederer, M. 2002. "Simultaneous interpretation: Units of meaning and other features." In *The Interpreting Studies Reader,* F. Pöchhacker and M. Shlesinger (eds), 131–140. London/New York: Routledge.

Lee, T.-H. 2002. "Ear voice span in English into Korean simultaneous interpretation." *Meta* 47 (4): 596–606.

———. 2003. "Tail-to-tail span: a new variable in conference interpreting research." *Forum: International Journal of Interpretation and Translation* 1 (1): 41–62.

Mackintosh, J. 1985. "The Kintsch and Van Dijk model of discourse comprehension and production applied to the interpretation process." *Meta* 30 (1): 37–43.

Moser-Mercer, B. 1985. "Screening potential interpreters." *Meta* 30 (1): 97–100.

Oléron, P. and Nanpon, H. 1965/2002. "Research into simultaneous translation." In *The Interpreting Studies Reader,* F. Pöchhacker and M. Shlesinger (eds), 42–50. London/New York: Routledge.

Paneth, E. 1957/2002. "An investigation into conference interpreting." In *The Interpreting Studies Reader,* F. Pöchhacker and M. Shlesinger (eds), 30–40. London/New York: Routledge.

Pöchhacker, F. 2004. *Introducing Interpreting Studies.* London/New York: Routledge.

Podhajská, K. 2007. "*Časový posuv v simultánním tlumočení z angličtiny do češtiny a jeho závislost na typu textu*" [*Time lag in simultaneous interpreting from English into Czech and its dependence on text type*]. (Master's thesis, Charles Univeristy, Prague.)

Podhajská, K. 2008. "Time lag in simultaneous interpretation from English into Czech and its dependence on text type." *Folia Translatologica* 10: 87–110.

Schilperoord, J. 1996. *It's about Time: Temporal Aspects of Cognitive Processes in Text Production.* Amsterdam: Rodopi.

Sharmin, S. Špakov, O. Räihä, K.-J. and Jakobsen, A. L. 2008. "Where on the screen do translation students look while translating, and for how long?" In *Looking at Eyes. Eye-tracking Studies of Reading and Translation Processing* [*Copenhagen Studies in Language* 36], S. Göpferich, A. L. Jakobsen and I. M. Mees (eds), 31–51. Copenhagen: Samfundslitteratur.

Shlesinger, M. 1998. "Interpreting as a cognitive process: What do we know about how it is done?." In *II Estudios sobre traducción e interpretación* [*Second Studies about Translation and Interpreting*], L. Félix Fernandez and E. Ortega Arjonilla (eds), 749–766. Málaga: University of Málaga.

Treisman, A. M. 1965. "The effects of redundancy and familiarity on translating and repeating back a foreign and a native language." *British Journal of Psychology* 56 (4): 369–379.

PART II

Process research in interpreting and translation

A new pair of glasses

Translation skills in secondary school*

Miriam Shlesinger*, # and Ruth Almog*
*Bar-Ilan University, Israel / #University of the Free State, RSA

Ever since 1989, students in some secondary schools in Israel have had the option of participating in the two-year Translation Skills Program, focused on written translation from English (L2) into Hebrew (L1), with particular emphasis on developing their meta-linguistic awareness while enriching their command of the two languages. Classes are taught by English teachers, who are first required to participate in a two-semester (112-hour) teacher training course focusing on translation skills, methods of teaching translation and the meta-linguistic insights one may glean from the process of translation. Throughout both courses – the one administered to the teachers and the one they then teach to the secondary school students – the point is made that the course is neither a foreign language (English) class, as such, nor a program for training full-fledged translators. Rather, it is a self-contained module aimed at enhancing students' grasp of inter-linguistic similarities and differences and exploring relations between language and culture. The paper reports on a study of the direct and indirect effects of the program on fourteen novices' development of meta-linguistic awareness. The study is based on both qualitative and quantitative data, including questionnaires, interviews and test results collected over the year during which the program was administered, and points out the advantages of introducing the translation skills program as an elective enrichment course at secondary school level and its potential contribution to language-related skills in general.

Keywords: secondary/high school students, written translation, meta-linguistic awareness, Hebrew, matriculation exam

---

* Birgitta Englund Dimitrova is an expert's expert, an authority on expertise, a scholar whose research has focused on the ascent towards a full understanding of the complexities of the translation process. The participants in the study described below, by contrast, have only now placed their foot on the bottom rung of the ladder and have no aspirations towards expertise (at this stage, at least). And so, the two ends of the continuum converge in a paper dedicated to Birgitta, about fourteen high school students experiencing their first encounter with the process of translation – which is also the process of learning, and of sharing insights, so that all of

1. Introduction

This paper is based on the collaboration of two authors. In 1989, the Chief Inspector for English in the Ministry of Education in Israel, Rafael Gefen, invited Miriam Shlesinger, the first author, to create and launch the syllabus for the new Translation Skills Program (TSP), to be introduced in grades 11 and 12 of Israeli secondary schools. The Ministry also undertook to subsidize the training course (112 hours) offered annually to prospective teachers, and later initiated the publication of a coursebook (Rosenbluth and Ballas 1998) as well. For nineteen years, Shlesinger was in charge of training the teachers, who went on to teach the TSP in their respective schools. The second author Ruth Almog, a graduate of the teacher training course, has taught the TSP to her own secondary school students for thirteen years. She is also a graduate of the Department of Translation and Interpreting Studies, and devoted her Master's thesis (under Shlesinger's supervision) to evaluating the effects of the TSP. The original research (Almog 2009) explored the progress of a group of secondary school students participating in a course in translation into their L1, and included analyses of their reading comprehension, their performance on written translation tasks and their levels of meta-linguistic awareness (MA). The present paper, based on that study, is confined to the latter. Its objective was to investigate students' awareness of the meta-linguistic issues posed by the translation task and by the differences between the two languages (and cultures) concerned. To do so, we focused our attention on the ways in which this awareness develops and manifests itself at the different stages of the translation skills program, as seen in the oral and written discourse elicited by the researcher from the fourteen student participants.

2. Translation skills as a self-contained course (the translation skills program)

To our knowledge, programs for teaching translation at secondary school level are quite rare; in fact, we know of no other full-fledged syllabus, besides the one described below.[1] From the start, the TSP was directed at students whose achievements in both

---

us, novices and experts alike, may end up just a little wiser. I am grateful for the privilege of knowing Birgitta, and of having this opportunity to pay tribute to her. (I am also grateful to my co-author for enabling me to dedicate our joint research to Birgitta.)

1. Clearly, this may be a lacuna on our part; i.e. such programs may exist, but we have found nothing about them in the Translation Studies literature. Translation as part of the foreign language curriculum is frequently discussed, but not as a separate entity or discrete module. Moreover, almost invariably, the students described in the literature are either undergraduate or (more frequently) graduate students. In other words, while the teaching of translation at

their L1 (Hebrew)[2] and their L2 (English) final exams had been high (grades no lower than 85%). The official course description spoke of a total of six yearly credits (3 hours per week in the 11th grade and another 3 hours per week in the 12th grade), but many schools opted for a leaner version, totaling 2 hours per week in each year (or, in some cases, compressing the course into a single year). The TSP is currently offered as an elective at about eighty secondary schools throughout the country. About 700 students a year opt to take the nationwide matriculation exam, but students are welcome to enroll in the program even if they choose not to do so. (Most have accumulated more than enough matriculation points for university admission purposes, and do not actually "need" the TSP exam.)

Describing the objectives of the course, the Inspector for English in the Ministry of Education wrote the following:

> […] the program includes, in addition to the techniques of translation, the following areas:
>
> – contrastive analysis of both language and culture
> – raising awareness of the way in which languages reflect cultures
> – awareness of the way language affects thinking within and across cultures
> – critical evaluation of resources.
>
> The mode of teaching is workshop oriented. Students work in pairs and in groups, reviewing each other's work, critically approaching different issues raised while translating and applying various higher-order thinking skills.
>
> (Dr. Judy Steiner, internal communication, 29 July 2008)

Class size ranges from 5 to 25 students, depending on the school, the threshold established by the administration, the number of students found eligible, and the number of those interested in taking an extra elective. The students' level of English proficiency is fairly homogeneous. Lessons are informal, with L1 and L2 being used interchangeably, and the ambience is relaxed, collaborative and non-competitive. Authentic texts are preferred to doctored or adapted ones, in the interest of providing a realistic, communicative task (cf. House 1981; Snell-Hornby 1985; Klaudy 1996; Stibbard 1998).

The TSP syllabus combines meta-linguistic discussions and the translation of short, authentic texts. In the process, students are exposed to a range of terms and concepts, presented by the teacher and supplemented by the course book (Rosenbluth

---

university level features prominently in the literature, we found no discussion of translation as a self-contained course at secondary school level.

2. In principle, the program was offered to speakers of Arabic as well; i.e. to students with Arabic as their L1, Hebrew as their L2 and English as their L3. Generally speaking, however, the program was very rarely adopted in Arabic-speaking schools. The study reported here focused on students with Hebrew as their L1.

and Ballas 1998).[3] Some of these are linguistic, par excellence, relating to the contrastive analysis of the two languages, while others are more closely aligned with the study of translation itself; among them are the following: the meaning of meaning, norms of translation, target audience, culture-specific items, semantic field, salient/relevant features, strategies, lexical voids, pragmatic voids, false cognates, translatability, reference materials, lexicalization, grammatical gridding (e.g. differences in the gender, tense, aspect etc.), register, dialect, genre, ambiguity, synonymy and polysemy,[4] collocation, connotation, homonymy, borrowings, neologisms, cohesion, coherence and sonority. Gradually they come to see the relevance of these concepts to the problems they encounter in translation, and the ways in which recurrent patterns of differences between the two languages and the two cultures affect translatability. Translation is thus seen not as an end in itself but as a means of gaining insights into the complexities of language and culture.

Beeby (2004) describes a syllabus that is similar in some ways to the TSP, albeit being aimed at undergraduates rather than secondary school students and – more importantly – having professional training as its ultimate goal. She notes that "declarative knowledge about contrastive rhetoric may be acquired in the first and second years, but may not become fully procedural until the translator has been working for several years" (2004: 43). Since the TSP does not aim at training translators – or at imparting knowledge that will necessarily become fully procedural – this limitation is of little relevance. For TSP purposes, declarative knowledge is an end in itself, with procedural knowledge as a fringe benefit. L2 (as well as L1) enrichment is another welcome by-product, but is not the main aim of the program. In short, the TSP is neither a means of foreign/second language teaching, as such, nor a form of translator training, but a self-contained form of intellectual enrichment. Like Englund Dimitrova (2005: 61), we speculate that "performing a translation task shares important cognitive processes with monolingual writing tasks, namely *planning, text generation* and *revision*," for which reason performing translation tasks in secondary school may, we believe, be conducive to better performance on tasks performed in the students' L1. We have not, however, conducted an empirical test of the possible benefits of the program for L1 writing skills.

As mentioned above, most translation courses taught today are designed either as a *means* of enriching foreign language learning among undergraduate students, or

---

3. To take a typical example, the first chapter (titled "Love at first sight") ends with the following brief summary: "In this unit you have translated your first sentences. Only when attempting to cope with them did you realize that translation is conveying meaning for meaning and not word for word. The translation of even simple and straightforward sentences is a process involving many considerations. You have also been introduced to difficulties of various types – grammatical and lexical difficulties, idiom and metaphor, and missing knowledge. In the next unit you will move from single sentences to a full passage. Good luck!"

4. This is but one of several "pairs" that pose a recurrent challenge in classroom discussions (Van Vaerenbergh 2007).

as an *end*; i.e. of preparing them for future employment (cf. Klein-Braley 1996; Sewell and Higgins 1996). When it comes to the former, the use of translation in the foreign language classroom is regaining its respectability, and is likelier to be a contextualized exercise rather than a study of words or sentences in isolation. (For an overview, see Kozminsky, Weizman and Horowitz 1998 and Schjoldager 2004). As for the latter, much of the literature on translator training is concerned with the pros and cons of alternative pedagogical approaches and the difficulty of determining "teachability of translation strategies," in the case of professional translators-to-be (González Davies, Scott-Tennent and Rodriguez Torra 2001: 737) and the effects of different forms of classroom intervention on the students' ultimate ability to perform as professionals (e.g. Shlesinger 1992; Kiraly 2000; González Davies, Scott-Tennent and Rodriguez Torra 2001; Englund Dimitrova 2002; Wilss 2004; González Davies 2005).

Notwithstanding the vocational orientation, there are those who also point to the contribution of translation courses to students' overall recognition of linguistic and cultural issues, independent of their professional benefits. Sewell (1996), for example, points to the collateral benefits of translation classes that have no vocational motivation and to their role in enhancing meta-linguistic awareness: "[…] because the exercise involves scrutinizing bits of language and evaluating equivalents in the target language, there is some increase in the students' mastery of both languages. […] The activity can increase verbal agility and awareness of the ways ideas can be expressed" (1996: 142–145). Describing a translation workshop as a language-learning tool for undergraduates, Conacher (1996) notes its usefulness in encouraging students to appreciate translation as an inroad into increased linguistic and cultural awareness and providing them with a motivating, challenging and enjoyable approach to language learning, and adds that "although the students may never become translators, they will never fail to appreciate the dedication, commitment and expertise of the professional translators they may one day employ" (1996: 180). Fraser (1996) stresses that "[…] process analysis not only sheds light on the features of successful professional practice but can also be applied to students learning to translate" (1996: 124). Bernardini (2004) too speaks of the importance of "providing learners with the awareness, resourcefulness and reflectiveness that they will certainly require" (2004: 27).

Unfortunately, none of the existing literature takes the students' age and level of schooling into account, although it stands to reason that a course administered at secondary school level will require methods and pedagogical strategies geared to this audience. After all, just as undergraduate students have specific needs that do not correspond to those of postgraduate students, who are usually older, more mature and more experienced (González Davies 2005: 71), so too the needs and expectations of secondary school students are fundamentally different from those of undergraduates.[5]

5. An analogy may be found in the study of MA in bilinguals: "Adult bilinguals, because of their more developed linguistic abilities and metalinguistic awareness, have a greater range of strategies available to them […]" (Malakoff and Hakuta 1991: 149).

Clearly, then, the syllabus and the pedagogical techniques should take into account age, maturity and life experience.

3. The study

The longitudinal study described here continues along the lines of Kozminsky, Weizman and Horowitz (1998), who investigated the effects of the TSP on students' meta-linguistic awareness, and found a marked improvement – compared to the results for a matched control group – as manifested in an "explanation task," in which participants were instructed to read a source text, to select five words or phrases that appeared difficult or problematic for translation and to analyze these. (In line with the rationale of the program, the TSP matriculation exam too includes a section – worth 30 points out of 100 – requiring explanations of translational problems and solutions.) The authors saw the results of their study as providing possible insights into the nature of secondary school students' meta-linguistic skills, and into the potential for enhancing those skills through the study of translation and the discussion of problem-solving strategies.

The present study was carried out by the second author in 2007–2008, and centered on the hypothesis that the study of translation skills (focusing on translation from L2 into L1) in secondary school, and the collaborative methods used in the translation classroom, would increase students' MA;[6] i.e. their "awareness of the underlying linguistic nature of language use [which] allows the individual to step back from the comprehension or production of an utterance in order to consider the linguistic form and structure underlying the meaning of the utterance" (Malakoff and Hakuta 1991: 147). This also refers to their ability to analyze language as a "thing," a "process" and a "system," to "think about a language" (Cazden 1974), to function objectively outside one language system and to objectify languages' rules, structures and functions (Mora 2001).

In the context of the present study, the TSP was taught from October 2007 until June 2008 (the students' first and only year in the program).[7] More often than not, students were divided into study groups of three or four, in which they checked

6. When he first proposed the idea of the TSP, Gefen (1982) perceived translation as "a practical skill that could be applied by the students after they graduated from high school […] useful if students needed to translate a letter, for instance," (p. 15) and perhaps through exposure to translation practice students "might even be able to develop some insight into language as such, precisely through the direct confrontation between the mother tongue and the foreign language" (p. 19). Meta-linguistic awareness was addressed not as an objective but as an additional benefit.

7. The school referred to here was among those that opted for a "leaner version," of the course totaling only one year, with two weekly hours throughout. While such an arrangement is not encouraged by the Ministry, the decision ultimately lies with the school administration.

homework, translated together or discussed issues that had arisen in the process of translating – with the teacher interacting with the groups. They were required to keep a log of their progress and of terms they had been taught, for the purpose of describing translational challenges, or differences between their two languages (cf. Cronin 2005, for a discussion of the role of terminology in allowing the theorist to stand back and describe what is happening, and of facilitating understanding.) They were also encouraged to reflect on translation solutions they regarded as particularly good. The process was collaborative, with students often debating the types and causes of shifts, as well as alternative solutions.[8] The study comprised four stages.

3.1 Stage I – Pre-TSP translation and questionnaire (October 2007)

To establish a baseline, the students were asked to translate a 180-word text (Appendix I) and to answer a questionnaire (Appendix II) reflecting on the process they had undergone while translating.

3.2 Stage II – Post-TSP translation and questionnaire (May 2008)

Towards the end of the program, students were asked to translate the same text they had translated in Stage I, and to describe the translation process, using the same questionnaire.

3.3 Stage III – Self-assessment questionnaire (June 2008)

A self-assessment questionnaire was administered, designed to provide a quantitative measure of students' own perception of developments in their MA.

3.4 Stage IV – Interviews (June 2008)

Individual interviews were conducted to elicit feedback and validate the questionnaire data.

The study was constrained by a number of methodological limitations:

1. Class attendance was not always regular, and levels of motivation, commitment and conscientiousness were uneven.
2. The study did not include a control group.

8. The course administered to prospective TSP teachers makes frequent recourse to the social-constructivist methods advocated by Kiraly (2000) and his emphasis on collaborative learning.

3. The study was confined to a single class of students (n = 14) in a single school, taught by a single teacher.
4. The teacher was also the researcher.

4. Findings

Most of the data presented below are verbal and qualitative, and as such, are open to multiple readings. Moreover, our selection of the excerpts from the students' oral and written feedback may reflect our own bias, although we have aimed for a balanced cross-section.

4.1 Stage I – Pre-TSP translation and questionnaire

When asked to describe the translation process during the first week of the course, none of the students described the process per se, which is not surprising, since they had not been made aware that such a process existed.

Their descriptions read as follows:[9]

- Student 1 and 7: "I read each line first; then I translated."
- Student 2: "I didn't know how to translate these words or phrases."
- Student 3: "I wasn't sure which word to use."
  - "I changed the sentence because it sounded odd."
  - "I thought a lot."
- Student 4: "I didn't translate word for word, but according to sentences."
- Student 5: [no answer]
- Student 6: "The sentence was unclear in English."
- Student 8: "I was frightened by the number of words I didn't know."
- Student 9: [no answer]
- Student 10: "I was under pressure."
- Student 11: "I was mostly thinking and looking up words."
- Student 12: "I was staring at the paper, mostly, wondering how to begin."
- Student 13: "I didn't know what to do with the 'slang'".
- Student 14 claimed that he didn't know what was meant by "translation process."

The students had been provided with reference materials (monolingual and bilingual dictionaries, and a bilingual thesaurus of idioms and phrases), to be used as they saw fit. Use of reference materials ranged from 3 to 15 times per student (mean: 10.25 times).

9. English translations of the students' comments are our own, throughout.

4.2 Stage II – Post-TSP translation and questionnaire

Not surprisingly, by the end of the year, the students were far more verbal, and made recourse to a rich array of descriptors, reflecting their perceptions of the translation process. Their answers – referring both to their understanding of the source text and to their formulation of the target text – were more nuanced and more detailed. Most began their report by noting that they had read the text first (sometimes twice), and then went on to discuss various specific challenges, and strategies for coping with these. Clearly, their formulations may have been affected by the fact that their teacher was to read these questionnaires; nevertheless, the difference between these reports and those of Stage I – as this concerns an awareness of lexis, register, etc. and their relevance to the target text – speaks for itself. We have selected eight of the most vivid and reflective excerpts, each relating to one of the themes that figured in the class discussions (register, slang, collocation, reference materials, voids, grammar, subtleties of meaning, etc.):

- First I read the text from beginning to end. If I didn't understand a word, I tried to understand it according to the text. If I still didn't understand it, I used the dictionary (Hebrew – English) to avoid "not knowing what I don't know." I deliberated over the register, because it changed from paragraph to paragraph.
- After finishing the translation, I decided to delete the slang terms for *cigarettes*, because in Hebrew the register works differently. I translated the title at the end, because I wanted to use an idiom, like in the source text, but I knew I had to complete the whole text so I could find a sensible equivalent idiom or some other collocation in Hebrew. The word "inserted" was difficult, because of the gridding, and I couldn't find the right word for "consumed" in Hebrew (a) because of the passive and (b) because of the gridding.
- I first read the whole text the way we were taught to, so I could get at the main idea as well as the register and the tone. I translated the text quickly, but marked the problematic words or phrases. I used the dictionary to help me with these passages. I was preoccupied by two major problems: The use of the passive in the source text, and the use of slang expressions for *cigarettes*. I know there are a lot of slang terms for drugs in Hebrew, but not for cigarettes. This is an interesting sociological fact in itself. I guess we could call it a sociological void, but I don't remember learning that particular type of void. Perhaps cultural? Anyway, I decided to omit them. I had a problem with the collocation "finely cut". I wasn't sure what it meant exactly. Then I attempted to get the register right – it kept changing throughout the text. Then I checked my grammar in Hebrew. Then I sat for 15 minutes to let the text "sit." Then I made the final adjustments and translated the title. I think it took me a lot longer this time.
- First I read the whole text and then I began translating. I marked the words I thought I'd have difficulty with, and then I looked them up in the dictionary. I stared at the blackboard for a while, deciding which words would fit the best. I got

stuck on the slang term for *cigarettes*, because there were no entries, and as far as I know there are no equivalent slang terms in Hebrew. The word "disease" confused me too, because according to the text the word really should be "problems." I had a problem with the title, which I translated last, because I wanted to find the right collocation. I even had trouble with subject-verb agreement. I realized that I had used the male form of the verb when I should have used the female, or vice versa.

– I read the text from beginning to end and began translating, leaving the title for last. I looked up all the words I didn't know, or wasn't 100% sure that I knew. I had a problem with the slang terms for *cigarettes*, and I admit – I left the classroom to see if I could find anyone outside who knew any slang terms in Hebrew for cigarettes. I even texted my friend. When I got to the title, I used the *Thesaurus of Idioms and Phrases* to find a suitable collocation. I had trouble with the words "finely," "commonly used," "cigarette holder." The word "devices" has a larger gridding in English. I don't think we have exactly that term in Hebrew. I shortened the sentences and changed a lot of them from passive to active.
– I read the text twice and looked up difficult words or expressions. When I began to translate, I read each sentence two or three times. I tried to understand the *meaning* of each sentence and not just the words. I tried to overcome the pitfalls (there were a lot). I split up the longer sentences. I changed a lot of sentences from passive to active. I had difficulty with time expressions, such as "generally," "usually," because when I translated them they didn't make sense in Hebrew. There were examples of ellipsis that confused me because I wasn't always sure what the verb was referring to. The word "commonly" is a morphological void and it was hard for me to find the equivalent for "commonly known." The word "colloquially" is both a morphological and a lexical void, and I had trouble finding an equivalent for "colloquially used."
– While translating, I was looking up words in dictionaries. I was concerned it would sound like translationese. I wasn't sure whether the words "finely cut" meant cut in little pieces or that it was cut well. There was no equivalent in Hebrew for the word "smolder" so it is a lexical void. I changed most of the passive sentences into active, because we don't use the passive as much in Hebrew. I divided the long sentence into two sentences to improve the Hebrew syntax.
– I found this text even harder the second time. I was more aware of the subtleties of the English. There was a lot of use of the passive voice in English that I had to change, but my Hebrew grammar isn't as good as it should be, I guess, because some of the sentences really didn't sound right. I used the dictionary a lot. I don't remember how many times.

The students had been provided with the same reference materials (monolingual and bilingual dictionaries, and a bilingual thesaurus of idioms and phrases) as in Stage I – Use of the reference materials ranged from 0 to 15 times (mean: 13 times).

Table 1. Students' responses to post-course questions referring to the TSP

Scale	Number of responses for each question (n = 14)								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	0	0	0	1	0	0	1	0	3
2	1	1	0	0	0	0	2	1	4
3	3	1	2	3	1	4	3	3	6
4	5	7	7	1	9	7	5	6	1
5	5	5	5	9	4	3	3	4	0
Mean score	4	4.1	4.2	4.2	4.2	3.9	3.5	3.9	2.4

4.3 Stage III – Self-assessment questionnaire

A separate questionnaire administered to the students at the end of the TSP focused on their own perception of changes in their command of translation skills and attitudes towards what they had learned. Using a Likert scale (1 = not at all; 5 = very much), students responded to the following descriptions of the process they had undergone (Table 1):

1. I am more aware of the types of errors I am likely to make when I translate.
2. I make fewer errors when translating.
3. I know where to find the vocabulary I need when translating.
4. I am more aware of the differences between English and Hebrew.
5. I understand why there are translation errors in the media.
6. I make more use of reference materials.
7. I will make use of my translation skills in the future.
8. I learned more than I thought I would in the course.
9. I would like to pursue a career in translating.

Overall, the students' self-assessment of their progress was high (mean: 4.0), particularly in the case of questions 1–5, which related directly to MA.

4.4 Stage IV – Interviews

All fourteen students were also interviewed by the second author, their TSP teacher. They were informed that their answers would be used for research, but were not told what the research would focus on. The phrasing of the questions aimed to be non-suggestive, out of consideration for the possibility of students being inclined to provide positive feedback, but the identity of the interviewer cannot be ruled out as a factor. The answers are largely reminiscent of those reported by Conacher (1996), whose undergraduate translation workshop sought to encourage students to appreciate translation as an inroad into increased linguistic and cultural awareness and

to provide them with a motivating, challenging and enjoyable approach to language learning. She points to marked improvements in grades and almost unanimously positive feedback.

> […] It does then seem to have been able to provide a motivating, challenging and enjoyable approach to learning. […] These students may never become translators, but they will never fail to appreciate the dedication, commitment and expertise of the professional translators they may one day employ. (1996: 180)

The answers below point to an overall favorable evaluation of the TSP; of course the students' ability (e.g. on questions 1 and 5) to reconstruct their expectations and impressions at the outset of the course was limited, and was presumably colored by their later experiences.

Question 1: *What did you expect to learn from the TSP?*
Ten of the students said they had expected to learn the differences between English and Hebrew, on both a grammatical and a lexical level, and how to bridge the gap between the two languages. Two students mentioned expecting to learn about the process of translation and how to translate more quickly and with fewer errors. They also mentioned that they hadn't realized translation was so "abstract" and was an acquired skill, although one must have a certain talent in order to do it well. Two students said that they hadn't given it much thought prior to the course but had felt the course would widen their horizons, so that no matter what, they would benefit. One student claimed she had heard the course was "fun," that it didn't entail a tremendous amount of extra homework, and that it had some "prestige value." She also maintained that she did not "need" the extra two matriculation points and had felt she could simply drop out if it did not meet her expectations.

Question 2: *In what ways did the curriculum meet – or not meet – your expectations?*
Only eleven students elaborated on this question. Among their statements:

- "I learned much more than I thought I would."
- "I would like to learn even more about the differences between the languages (Hebrew and English) and more about the history of the languages."
- "I was surprised to learn things I hadn't thought about before, like about the cultural voids between the languages, that cannot always be overcome in translation and about the difficult choices that the translator needs to make in order to find the 'mot juste.'"
- "I'm sorry the course was so short. I think it should have started in eighth grade."
- "I learned more than I had expected. I thought we would learn the 'right' way to translate, but then on second thought, we learned there really is no 'right' way – that's what makes it so hard – or easy."

Question 3: *What TSP subjects may prove relevant in real-life situations?*
One student mentioned a conversation with a friend, in which the friend kept using the word "sad" to describe something. The student told her friend that she didn't really mean "sad," but something else. She opened the thesaurus and found the word "upset." She said, "That was it! That was the word I was looking for. Before the course, I would not have realized the subtle differences between words, and I certainly wouldn't have taken the time to look them up in a thesaurus – I didn't even know what a thesaurus was! It may seem silly to most people, but it's a great feeling when you are able to find the right word to express yourself – in any language." Another student said it would take some time before she could read for pleasure again. She found herself back-translating every translated book she read, and analyzing why she found a passage awkward. She added that she felt more confident when criticizing translated literature because she had more of an understanding and the means to express herself. Three students stated that unless they were going to pursue a career in translation, the skills they had learned were not relevant to their daily lives – but all three added hopefully that anything they learned in life could be useful, and they were sure they would see this later. Nine of the students said they already felt that the skills they had learned were relevant to the way they spoke, read literature and wrote. One of them mentioned that learning about pragmatics had made him more aware when he wrote in both Hebrew and English. He found himself listening closely to the way people speak, and analyzing intensifiers and hedges. One student said she had never paid much attention to subtitles, but now she could understand why there were so many mistakes.

Question 4: *Did the course affect your English (L2) reading comprehension? If so, how?*
Some of the students related not only to the question of reading comprehension in their L2, but also to the effect of the course on their ability to express themselves well in their L1 (even though this was not part of the question). Twelve of the students said that they believed they had a better ability to comprehend texts in English. Five of them stated that perhaps their immediate comprehension was not better, but they were aware of how to approach a text properly in order to understand what it really meant. Two students claimed that the course had had no effect on their Hebrew or their English. The others mentioned specific aspects of both languages that were affected by the course: their appreciation of translated literature, their awareness of how difficult it was to translate a text, or the amount of time one could spend deliberating over a single word. They said they had been made aware that "synonyms are never really synonymous," and that spoken Hebrew was a relatively "new" language, borrowing so many words from other languages, especially English – which made them aware of false cognates. Another student mentioned a tendency to use more metaphors in writing and even more when speaking. She had begun to pay more attention to the "beautiful Biblical expressions we have in Hebrew," and even looked up their origins, while before the course she had taken them for granted. She said it had never occurred to her that there were reference materials to use for this purpose. Three

students stressed becoming more aware of the richness of English and of the many voids in both languages in relation to the other. One student mentioned paying more attention to the differences between the languages and was making more of an effort to speak correctly. She said that, ironically, she noticed this because her friends said she sounded "different" when she spoke. She said she was using less slang than she used to. Several students said they were more aware of translationese.

Question 5: *How is the process of translating the text different now?*
Most of the students said they could not remember how they had felt about the translation process at the beginning of the program, speculating that perhaps this was due to the fact that they had not thought of it as a process. Two students said it took them longer to translate now because they were more aware of the different items they had to think about. They were making more use of reference materials, and were less daunted by the task.

Question 6: *What did you actually learn in the TSP? What will you take with you?*
Although many of the students related to the same aspects of the course, each of them also referred to a number of issues that they regarded as particularly valuable. The following are (our translations of) their complete answers to this question.

Student 1: I notice bad translations in the media. Even if it's not really a mistake, you know, like a non-binary error, but the sentence just doesn't seem right.

Student 2: I notice specific words in English and I try to figure out their origin, especially idioms and metaphors, so if I were to translate them, I would know exactly what they mean, and then it would be easier for me to find an equivalent in Hebrew. I also notice the different tenses in English more, especially the perfect [sic] and differences in grammatical nuances between Hebrew and English.

Student 3: I mostly notice translations more when I read. It never occurred to me to find out who translated a book, or to give the translators credit for their hard work. By the way, their name should be on the front of the book, right along with the author. I also learned all sorts of terms I'd never heard before. It's opened up a whole new world for me.

Student 4: I pay more attention to the way I talk. I must say – it can be very frustrating. I didn't used to pay attention to gender in Hebrew. Now I find myself asking my friends, "Wait – is such and such masculine or feminine?" I'm slowly losing my friends [giggle].

Student 5: I've become much more pedantic about my use of language, both in Hebrew and in English, even though before the course I also did my best to write properly (or so I thought), but now I'm even more careful.

Student 6: I notice things like false cognates. I don't think I really know yet what I will "take with me" because the year has just ended. Maybe if you ask me next year I will have a more honest answer for you.

Student 7: I always thought I knew Hebrew really well. I mean, it's my mother tongue. But now I know that I had (and have) a lot to learn. I won't even get into what I still need to learn in English!

Student 8: I pay attention to everything relevant to adjectives, tenses, gridding, vocabulary, idioms, syntax or punctuation. I learned so much; I can't even remember it all.

Student 9: I learned how cohesive devices can change the whole meaning of a sentence, I mean conjunctions, like "although," or "but," or "in spite of the fact" … I learned new expressions, and loads of translation terms that I'd never even heard of. I learned about the subtleties of the past tense in English that we don't have in Hebrew. I know I was supposed to have learned all this in the lower grades, but for some reason, in the translation course it became much clearer.

Student 10: I learned the definitions of different problems that arise in translation, like idioms, for instance.

Student 11: I became aware of different types of errors and the reasons behind them. I never thought I'd be able to notice errors or translationese in subtitles or advertisements.

Student 12[10]: I am really sorry, but I can't honestly say what I learned. I missed a lot of classes, and I really was so bogged down with my other subjects. I'm sure I could have gotten more out of the course if I had applied myself, but you know me … I think that if I ever attempted to translate anything I wouldn't feel very sure of myself.

Student 13: Throughout the course I had the chance not only to translate, but also to learn about the topics of my translations and about both languages. I mean, we learned so much stuff!

Student 14: I learned all kinds of expressions and grammatical structures such as voids, gridding, et cetera.

5. Discussion

It has been found that children who have a more developed MA are also likely to have more developed language skills in general (Malakoff and Hakuta 1991: 148). The three forms of elicitation used in this study point to an increase in students' MA and in their ability to articulate their own translation experience, using meta-linguistic and meta-translational terms and concepts. While the role of MA in the (bilingual) context of translation is most often discussed in the context of (professional) translator training – particularly as this concerns the bolstering of self-confidence (Kussmaul 1995: 31) – it may also play a vital role in a non-vocational framework, such as the one described here. In fact, many of those who write about translation as

10. This student's attendance and active participation had been weak and sporadic.

part of vocational training are not oblivious to these collateral benefits: "The importance of fostering in students a sharpened awareness of the thinking processes leading to their translations can never be over-estimated" (Sin 2002: 43). As Sewell (1996) has put it in connection with an undergraduate course:

> [...] The translation course begins to look like a course in text-awareness. But why not? Text-awareness underpins the professional translator's competence, but it also puts the language learner in a better position to make appropriate choices about TT phrasing. (1996: 145)

A student in the translation course run by Snell-Hornby (1985) expressed a similar sentiment:

> A translation course which seriously analyzes the possible choices of words for transferring meaning from one language to another is an invaluable tributary to the understanding of the linguistic structures of both languages. I find such courses enormously helpful. (1985: 26)

Translational proficiency might be thought of as an interplay between bilingual proficiency and meta-linguistic maturity, involving the recognition of commonalities and differences in the nature and functions of languages, analysis of linguistic knowledge and control over processing. It is both an awareness and a skill, as it rests on the ability to successfully approach and solve certain problems (Bialystok 1991; Malakoff and Hakuta 1991).

In the case of the students who participated in this study, levels of meta-linguistic awareness were clearly shown to have improved. While we also assume that their linguistic skills – reading comprehension, written expression etc. – in the two languages were favorably affected, the present study has not covered these. We have, however, tried to establish that the TSP played a significant role in the development of students' meta-linguistic awareness. As their responses to the self-assessment questionnaire (see Stage III – especially the answers to question 4) and their own statements (see Stage IV) seem to indicate, their subjective evaluation of the course points to a heightened awareness of language issues, in general, and of their own progress, in particular.

6. Conclusions

As noted above, our aim in the study reported here was to establish whether a secondary school course devoted to the teaching of written translation may enhance students' awareness of the particularities of their respective languages in relation to each other i.e. to increase their meta-linguistic awareness. Of course, like many studies dealing with the process of translation, it represents no more than a small number of subjects over a relatively short time. In keeping with the findings of Kozminsky, Weizman and Horowitz (1998), it points to possible benefits of using translation as a (discrete)

complement to regular EFL classes at the upper levels of secondary school. After more than two decades, the TSP seems to be a permanent fixture in the Israeli secondary school curriculum, albeit on a very small scale.[11] Students' self assessments – and teachers' feedback – indicate that most of them see the course as highly beneficial, and both quantitative and qualitative data provide evidence of its positive effects e.g. students' realization that language transcends the level of word-for-word equivalence, and entails cultural, sociological and historical aspects.

Meta-linguistic awareness is not only a *precondition* for becoming an expert professional translator; it is also a *by-product* of studying translation, whether taught as a self-contained program or as part of the foreign language syllabus. When it comes to the meta-language for describing the translation process and reflecting on one's own learning and translation strategies (cf. Chesterman 1998), the kind of MA fostered by collaborative and open discussions typical of the TSP are seen as highly beneficial to the intellectual development of secondary school students.

Despite the usual administrative, logistical and financial obstacles since its inception in 1989, the TSP has continued uninterrupted. In principle, it could easily be explored in other countries and in other language combinations as well – offering secondary school students elsewhere an enjoyable and useful addition to their regular curriculum. Indeed, as one of the students put it at the end of the course: "It is as if I have a new pair of glasses. I see things differently."

References

Almog, R. 2009. "Teaching the translation skills program in an Israeli high school and its effect on reading comprehension skills in L2 and metalinguistic awareness." (Master's thesis. Department of Translation and Interpreting Studies, Bar Ilan University, Ramat Gan, Israel.)

Beeby, A. 2004. "Language learning for translators: Designing a syllabus." In *Translation in Undergraduate Degree Programs,* K. Malmkjaer (ed.), 39–65. Amsterdam/Philadelphia: John Benjamins.

Bernardini, S. 2004. "The theory behind the practice: Translator training or translator education?" In *Translation in Undergraduate Degree Programs,* K. Malmkjaer (ed.), 17–29. Amsterdam/Philadelphia: John Benjamins.

Bialystok, E. 1991. "Metalinguistic dimensions of bilingual language proficiency." In *Language Processing and Language Awareness by Bilingual Children*, E. Bialystok (ed.), 113–138. Cambridge, UK: Cambridge University Press.

Cazden, C. 1974. "Play and metalinguistic awareness: One dimension of language experience." *Urban Review* 7 (1): 78–79.

11. Unfortunately, the status of the TSP as an elective tends to limit enrollment; principals often see it as a "frill," and are reluctant to allocate the extra teaching hours. English teachers, on the other hand, vie for the privilege of teaching it; after all, the course is offered to the best students and adds variety to the teacher's standard fare.

Chesterman, A. 1998. "Communication strategies, learning strategies and translation strategies." In *Translation and Language Teaching: Language Teaching and Translation,* K. Malmkjaer (ed.), 135–144. Manchester, UK: St. Jerome Publishing.

Conacher, J. E. 1996. "Native speaker to native speaker: Crossing the translation gap." In *Teaching Translation in Universities: Present and Future Perspectives*, P. Sewell and I. Higgins (eds), 161–182. London: CILT.

Cronin, M. 2005. "Deschooling translation." In *Training for the New Millennium: Pedagogies for Translation and Interpreting,* M. Tennent (ed.), 249–265. Amsterdam/Philadelphia: John Benjamins.

Englund Dimitrova, B. 2002. "Training and educating the trainers – a key issue in translators' training." In *Teaching Translation and Interpreting 4,* E. Hung (ed.), 73–82. Amsterdam/Philadelphia: John Benjamins.

———. 2005. *Expertise and Explicitation in the Translation Process.* Amsterdam/Philadelphia: John Benjamins.

Fraser, J. 1996. "'I understand the French, but I don't know how to put it into English': Developing undergraduates' awareness of and confidence in the translation process." In *Teaching Translation in Universities: Present and Future Perspectives*, P. Sewell and I. Higgins (eds), 121–134. London: CILT.

Gefen, R. 1982. "Introducing the skill of translation as one of the skills taught in secondary school (grades 10–12) and tested in the *Bagrut* examination." *English Teacher's Journal* (Israel) 27: 11–14.

González Davies, M. 2005. "Minding the process, improving the product: Alternatives to traditional translator training." In *Training for the New Millennium: Pedagogies for Translation and Interpreting*, M. Tennent (ed.), 67–82. Amsterdam/Philadelphia: John Benjamins.

González Davies, M., Scott-Tennent, C. and Rodriguez Torra, F. 2001. "Training in the application of translation strategies for undergraduate scientific translation students." *Meta* 46 (4): 737–744.

House, J. 1981. *A Model for Translation Quality Assessment*. Tübingen: Narr.

Kiraly, D. 2000. *A Social Constructivist Approach to Translator Education: Empowerment from Theory to Practice.* Manchester: St. Jerome Publishing.

Klein-Braley, C. 1996. "Teaching translation: A brief for the future." In *Teaching Translation in Universities: Present and Future Perspectives*, P. Sewell and I. Higgins (eds), 15–30. London: CILT.

Klaudy, K. 1996. "Quality assessment in school vs. professional translation." In *Teaching Translation and Interpreting 3,* C. Dollerup and V. Appel (eds), 197–204. Amsterdam/Philadelphia: John Benjamins.

Kozminsky, E., Weizman, E. and Horowitz, H. 1998. "Teaching translation in Israeli high schools." *Meta* 43 (1): 119–129.

Kussmaul, P. 1995. *Training the Translator.* Amsterdam/Philadelphia: John Benjamins.

Malakoff, M. and Hakuta, K. 1991. "Translation skill and metalinguistic awareness in bilinguals." In *Language Processing and Language Awareness by Bilingual Children*, E. Bialystok (ed.), 141–166. Oxford: Oxford University Press.

Mora, J. K. 2001. *Metalinguistic Awareness as Defined Through Research*. http://coe.sdsu.edu/people/jmora/MoraModules/MetaLingResearch.htm. (Accessed March 2010.)

Rosenbluth, P. and Ballas, N. 1998. *Two-way Traffic – Teachers' Guide and Pupils' Book.* Jerusalem: Ministry of Education.

Schjoldager, A. 2004. "Are L2 learners more prone to err when they translate?" In *Translation in Undergraduate Degree Programs,* K. Malmkjaer (ed.), 127–149. Amsterdam/Philadelphia: John Benjamins.

Sewell, P. 1996. "Translation in the Curriculum." In *Teaching Translation in Universities: Present and Future Perspectives*, P. Sewell and I. Higgins (eds), 135–159. London: CILT.

Sewell, P. and Higgins, I. 1996. "Introduction." In *Teaching Translation in Universities: Present and Future Perspectives*, P. Sewell and I. Higgins (eds), 9–14. London: CILT.

Shlesinger, M. 1992. "Lexicalization in translation: An empirical study of students' progress." In *Teaching Translation and Interpreting: Training, Talent and Experience*, C. Dollerup and A. Loddegaard (eds), 123–127. Amsterdam/Philadelphia: John Benjamins.

Sin, K.-K. 2002. "Myths and misconceptions in translation teaching." In *Teaching Translation and Interpreting 4,* E. Hung (ed.), 31–43. Amsterdam/Philadelphia: John Benjamins.

Snell-Hornby, M. 1985. "Translation as a means of integrating language teaching and linguistics." In *Translation in Foreign Language Teaching and Testing,* C. Titford and A. E. Hieke (eds), 21–28.Tübingen: Narr.

Steiner, J. 2008. "Proposal for including the subject of translation as part of the Pedagogical Horizons for Learning program". Internal communication by the Inspector for English, Ministry of Education, Jerusalem.

Stibbard, R. 1998. "The principled use of oral translation in foreign language teaching." In *Translation and Language Teaching: Language Teaching and Translation,* K. Malmkjaer (ed.), 69–76. Manchester: St. Jerome Publishing.

Van Vaerenbergh, L. 2007. "Polysemy and synonymy: Their management in translation studies dictionaries and in translator training. A case study." *Target* 19 (2): 235–254.

Wilss, W. 2004. "Translation studies: A didactic approach." In *Translation in Undergraduate Degree Programs,* K. Malmkjaer (ed.), 9–15. Amsterdam/Philadelphia: John Benjamins.

Appendix I

Text for translation

Please read the questionnaire **before** you begin the translation.
Please note the time you begin your translation (including reading time): _____
Circle words and phrases that were difficult for you to translate.
Please do not forget to translate the title.

Up in Smoke

A cigarette is a product consumed via smoking and manufactured out of cured and finely cut tobacco leaves, which are combined with other additives, then rolled or stuffed into a paper-wrapped cylinder (generally less than 120 mm in length and 10 mm in diameter). The cigarette is ignited at one end and allowed to smoulder for the purpose of inhalation of its smoke from the other (usually filtered) end, which is usually inserted in the mouth. They are sometimes smoked with a cigarette holder. The term cigarette, as commonly used, refers to a tobacco cigarette but can apply to similar devices containing other herbs, such as cannabis. They are colloquially known as 'cigs', 'smokes', 'ciggies', 'cancer sticks', 'death sticks', 'coffin nails' and 'fags'.

Cigarettes are proven to be highly addictive, as well as a cause of multiple types of cancer, heart disease, respiratory disease, circulatory disease and birth defects.

A cigarette is distinguished from a cigar by its smaller size, use of processed leaf, and white paper wrapping. Cigars are typically composed entirely of whole leaf tobacco.

* * * * * *

Please note the time you finished your translation completely and were ready to hand it in: _____

Appendix II

Questions about the translation process

Dear Students,
The following information will be used for a thesis in Translation Studies. Your careful answers will be very helpful in attaining insights into the process of teaching and learning translation skills in high school. Please answer the questions honestly, and fully. This will absolutely not affect your grade in any way. Your translation, as well as this questionnaire will remain anonymous. Thank you very much for your cooperation.

Personal information:

1. Age _____
2. Grade: _____
3. Gender:_____
4. English proficiency (circle): mother tongue / mother tongue level / excellent / very good
5. Hebrew proficiency (circle): mother tongue / mother tongue level / excellent / very good
6. I read in English for pleasure: yes _____ no _____

Questions related to your translation of the text

1. Please describe the process you went through while translating the text. Record a detailed account of what you did throughout the time used in translating the article (thinking, writing, looking words up, eating, drinking, staring, etc.).
2. Write down the term(s) that posed difficulty for you. ("Difficult" means that you were stuck, confused or that you deliberated for more than a few minutes, excluding help from the dictionary.)
3. What sentences did you revise? Please write the original sentence, and the revision. What was your rationale?
4. How many times did you use the dictionary?
5. What dictionaries did you use?
6. How many times did you use the thesaurus?
7. How long did it take you to translate this passage? ___ hour(s) ___ minutes

Are primary conceptual metaphors easier to understand than complex conceptual metaphors?

An investigation of the cognitive processes in metaphor comprehension*

Antin Fougner Rydning and Christian Lachaud
University of Oslo, Norway

The present behavioural study explores the cognitive processing of primary and complex conceptual metaphors during the first step of the translational process: comprehension. According to Lakoff and Johnson (1980), primary and complex conceptual metaphors are based on different conceptual architectures, a simpler one for primary conceptual metaphors, and a more sophisticated one for complex conceptual metaphors. As far as we know, there is no evidence that understanding primary conceptual metaphors requires less cognitive effort than understanding complex conceptual metaphors. In order to know whether primary conceptual metaphors are easier to understand than complex conceptual metaphors, a psycholinguistic experiment was designed and run to investigate processing speed and conceptual clarity during comprehension. Among other interesting results, our study provides evidence that conceptual clarity is greater with primary conceptual metaphors than with complex conceptual metaphors. Therefore, a translator will not be challenged in the same way depending on the type of metaphor s/he has to understand. As a first step towards understanding the mechanisms of comprehension, this study paves the way for further

---

* This volume to celebrate both Birgitta's birthday and her work is a welcome opportunity to touch base with her again, after several years of intermittent collaboration through the EXPERTISE team funded by the Norwegian Research Council. There came a time when we went our separate ways, as Birgitta decided to channel her energy into rounding up her pioneering book *Expertise and Explicitation in the Translation Process* (2005) while I decided to board the EYE-to-IT project, but I have continued to follow her work from afar. This volume is an apt opportunity for me to join up with colleagues who, like me, have benefitted from Birgitta's invaluable input, and to present her with some of my own research as a token of my respect for her work.
Antin Fougner Rydning

investigations on how comprehension influences the subsequent stages of translation, especially regarding how conceptual clarity may vary during comprehension and help determine the quality of the translation.

Keywords: primary conceptual metaphors, complex conceptual metaphors, conceptual clarity, comprehension

1. Introduction

Meaning construction is the focus of translational activity and creativity. Metaphorization is a central manifestation of meaning construction. In line with cognitive linguistics, metaphorization is viewed as a phenomenon of thought as well as of language.

According to the Contemporary Theory of Metaphor (Lakoff and Johnson 1980; Lakoff and Turner 1989; Lakoff 1993), our conceptual system is largely metaphorical, and conceptual metaphors are one of the central components of meaning construction. In order to understand the meaning of an abstract metaphorical expression, one first needs to conceptualize the underlying conceptual metaphor. Most of the time, this is performed automatically and effortlessly, below the level of consciousness (Lakoff and Johnson 1999). Remarkably, in everyday communication as well as during translation, processing metaphorical meaning seems as straightforward and natural as processing literal meaning. The following examples illustrate how this appears to be true:

(1) Figurative expression: This light is very harsh.

(2) Literal counterpart: This light is very intense.

(3) Figurative expression: His telephone is dead.

(4) Literal counterpart: His telephone is not functioning.

However, this subjective feeling may not reflect possible differences existing in the underlying cognitive processes. During comprehension, such differences may affect both the speed of understanding and conceptual clarity. In turn, variations in conceptual clarity may impact on the ultimate quality of the translation.

Grady (1997) and Lakoff and Johnson (1999) have suggested that conceptual metaphors be organized in two categories: primary and complex metaphors (as defined in the following section). Questions which arise are whether primary and complex conceptual metaphors are processed differently, and whether understanding a complex conceptual metaphor such as (3) induces higher cognitive efforts than understanding a primary conceptual metaphor such as (1). In other words, does easiness of conceptualization, which is likely to influence conceptual clarity, go hand in hand with the degree of conceptual complexity? To our knowledge, no one has yet probed cognitive processing differences between primary and complex conceptual

metaphors to answer these questions. It is the purpose of the present experimental study to (i) focus on the mechanisms of comprehension by introducing the factor of conceptual metaphors, and (ii) contextualise such mechanisms within the general framework of research on translation processes.

1.1 Primary vs. complex conceptual metaphors

Conceptual metaphors are ways of thinking about one conceptual domain in terms of another. The meaning of a linguistic expression such as "Birgitta is a warm person" can be understood as the mapping of the physical source domain of temperature onto the target domain of affection through repeated physical experience, i.e. feeling the warmth of another person's body when being held affectionately. In conceptual terms, the linguistic expression can be represented as AFFECTION IS WARMTH.

Johnson (1997) has pointed out that when subjective experiences are regularly conflated with sensorimotor experiences, associations are automatically built up between their corresponding mental representations. These automatic conceptual associations help structure the way we think.

1.1.1 *Primary conceptual metaphors*

A. *Sensorial universal aspects*

This "theory of conflation" has been further developed by Grady (1997), who sees that regular conflations of subjectivity and bodily experiences lead to the formation of primary conceptual metaphors. Primary metaphors arise naturally, automatically, and unconsciously through everyday experience by means of conflation, during which cross-domain associations are formed. Primary metaphors are foundational as they associate basic concepts through our perception and experience. Because such metaphors are directly grounded in the individual's sensorimotor experience of the world, they are likely to be shared by all humans. Consequently, they are expected to reflect ways of conceiving the world, manifesting in rather similar linguistic realizations across languages. For example, the linguistic utterance "Birgitta is a warm person," corresponding to the primary conceptual metaphor AFFECTION IS WARMTH, is expressed similarly in Norwegian (*Birgitta er en varm person*) and French (*Birgitta est une personne chaleureuse*). The universal nature of this primary conceptual metaphor could explain why its linguistic realizations in the three languages are the same.

B. *Influence from the physical environment*

Primary metaphors can also vary with the physical environment in which individuals live. These variations are reflected at the conceptual level, resulting in metaphorical expressions which are language and culture specific. Since Norway is a mountainous country, for instance, metaphorical expressions constructed with "mountains" are quite numerous, e.g. *Prosjektet står fjellstøtt* (back-translation: "The project stands mountain-steady," corresponding to "The project is solid as a rock").

Such specificities can also arise from the activities that environmental characteristics impose on the inhabitants of a particular country. In Norway, for instance, the popular sport of ski jumping has given rise to a number of ski jumping metaphors. Consider the following expression uttered by the director of a Norwegian oil company after the Norwegian Ministry of Petroleum and Energy had decided to construct a European CO2 test centre at the Mongstad power and heating station:

(5) […] *Mongstad vil koste mye og kreve mye. Men gleden er at vi vil lande i bunnen av bakken* […] (Norwegian Petroleum Society, January 20, 2010)
Back-translation: "Mongstad will cost a lot and require a lot. But the joy is that we will land at the bottom of the hill […]"

The conceptual metaphor underlying the metaphorical expression in (5) can be represented as

(6) SUCCESS IS A LONG AND SECURE SKI JUMP

It goes without saying that in countries where ski jumping is less popular, the idea of success will be conceptualized through other source domains, resulting in different metaphorical expressions. In British English a possible equivalent metaphorical expression could be based on the sport of cricket (e.g. "hit a six"), in which case the underlying conceptual metaphor would be

(7) SUCCESS IS GETTING THE MAXIMUM POINTS IN CRICKET

C. *Cultural specificities*
Different conceptual primary metaphors can also appear across cultures in order to express the same idea, leading to language specific varieties. Kövecses, who cites Ning Yu, suggests that "the more introverted character" of the Chinese people (2005: 70) reflects why the Chinese, in addition to sharing with English all the metaphorical source domains for happiness, also have the additional source domain OPENING. This conceptual primary metaphor underlies metaphorical expressions such as *xinhua nufang*, which translates directly as "flowers in the heart burst into bloom," and *kai xin*, meaning "open (up the) heart" (Liu 2010). In our Western societies, where we are supposedly more extroverted, our feeling of being happy is usually associated with being off the ground, as found in metaphorical expressions such as "I am in heaven" and "I am on top of the world." The underlying primary conceptual metaphor here is HAPPINESS IS UP.

We observe that a common target domain can be mapped onto different source domains due to environmental and cultural specificities, leading to dissimilar linguistic realizations.

1.1.2 *Complex conceptual metaphors*

Complex conceptual metaphors are based on the combination of primary conceptual metaphors through integrative processes, such as conceptual blending (Fauconnier

and Turner 2002). They consist in recombining the conceptual domains of primary conceptual metaphors into new blends, including additional forms of commonplace knowledge such as cultural models or folk theories. Because they can integrate culture-specific components, the same complex conceptual metaphors can result in different linguistic formulations across languages. Consider the following example:

(8) *Birgitta stuck to her course.*

This metaphorical expression corresponds to the complex metaphor A PURPOSEFUL LIFE IS A JOURNEY, built up by combining the primary metaphors PURPOSE IS DESTINATION, ACTION IS MOTION, and A MEANS OF ACHIEVEMENT IS AN ITINERARY, and relying on the cultural belief that humans have purposes in life and act in order to reach those purposes. When the concept of purpose is mapped onto that of destination, the concept of action is mapped onto that of motion, and the concept of means of achievement is mapped onto that of itinerary, we understand that Birgitta reached her initial goals.

In Norwegian, a fully corresponding linguistic expression does not exist. Among the many possible ways of expressing the same idea, one could e.g. say

(9) *Birgitta fulgte løypa hun hadde staket ut.*
Back-translation: "Birgitta followed the track she had marked out."

The complex metaphor is the same as in (8) above (A PURPOSEFUL LIFE IS A JOURNEY), as are two of the underlying conceptual primary metaphors (PURPOSE IS DESTINATION and ACTION IS MOTION). However, the third conceptual primary metaphor (A MEANS OF ACHIEVEMENT IS A MARKED TRACK), although retaining the same target domain (A MEANS OF ACHIEVEMENT), has a different source domain (A MARKED TRACK), which conveys the idea that the course is a track one has marked – a situation quite typical for Norway's cross-country skiers.

In French, a possible equivalent translation would be

(10) *Birgitta a gardé le cap.*
Back-translation: "Birgitta kept the cape."

Once again it is the third conceptual primary metaphor that differs, with A MEANS OF ACHIEVEMENT IS A TARGETED CAPE stemming from a different source domain, namely sailing (A TARGETED CAPE).

Although the meaning expressed in (8), (9) and (10) above is constant, the linguistic realization differs in the three languages. This cross-linguistic variation originates in the specificity of the source domain, while the target domain remains the same:

(11) A MEANS OF ACHIEVEMENT IS **AN ITINERARY** (English)

(12) A MEANS OF ACHIEVEMENT IS **A MARKED TRACK** (Norwegian)

(13) A MEANS OF ACHIEVEMENT IS **A TARGETED CAPE** (French)

Because complex conceptual metaphors may show greater cross-linguistic and transcultural variations, translating them may be more challenging for translators than translating primary conceptual metaphors.

2. Assumptions and hypotheses of the study

Traditional linguistics (e.g. Grice 1975; Searle 1979) contends that literal expressions are easier to understand than metaphorical ones. In his psycholinguistic studies, however, Gibbs (1990, 1994) found no processing speed differences between metaphorical and literal utterances, thus invalidating the traditional linguistic hypothesis and Glucksberg (2003) demonstrated that readers understand novel metaphors as quickly as their literal counterparts.

Where primary and complex conceptual metaphors are concerned, the possible similarities or differences in understanding them have hitherto not been studied. To our knowledge, only one experiment (Martikainen 2007) used translation as a method to test the psychological reality of these two kinds of conceptual metaphors, albeit following a different aim than ours. Because primary and complex conceptual metaphors rely on different conceptual architectures, their mental processing is likely to be different. Primary conceptual metaphors are assumed to be embodied, i.e. directly grounded in our everyday bodily experiences. Therefore, their comprehension is likely to occur through specific direct wiring of neurons in the brain (Narayanan 1997), between body-related brain areas and other brain areas involved in higher cognitive functions. Such neural structures would allow a fast, direct and easy processing of embodied meaning. Complex conceptual metaphors which do not have an experiential grounding of their own, but are grounded through their primary metaphor components, are supposed to require a more elaborated and indirect processing. In other words: the simpler the conceptual structure, the easier the comprehension, and probably, the greater the resulting conceptual clarity. The assumption that a metaphor's ease of comprehension is linked to its conceptual structure, however, contradicts our subjective experience of understanding both primary and complex conceptual metaphors as easily and naturally as literal expressions. Since this issue has not been studied extensively, and since it is important for highlighting the potential difficulties a translator might face during the initial stage of translation (conceptualization and comprehension), we intend to test whether figurative meaning in primary and conceptual metaphors are as easy to process as literal meaning.

3. Method

3.1 Independent variables

We manipulated the following four independent variables in the experiment:

i. Target metaphoricity (two levels: literal, figurative)
ii. Target complexity (two levels: primary, complex)
iii. Probe correspondence (two levels: yes, no)
iv. Task of the participant (two levels: detecting positive probe correspondence, detecting negative probe correspondence).

The first two variables were crossed according to a factorial design, generating four experimental conditions in the linguistic materials:

v. literal meaning obtained with the words used for creating primary metaphors
vi. literal meaning obtained with the words used to create complex metaphors
vii. figurative meaning obtained with the words used to create primary metaphors
viii. figurative meaning obtained with the words used to create complex metaphors.

The distinction between conditions (i) and (ii) is purely methodological, with these two conditions respectively serving as control for conditions (iii) and (iv). Indeed, literal expressions are not supposed to be primary or complex, according to cognitive linguistics. The last two variables are technical factors manipulated for the purpose of the task. The third factor allows us to create a situation in which the probe word carries a meaning which is either similar or dissimilar to that of the sentence. The fourth factor determines which task a participant has to do: answering in case the probe correspondence is either positive or negative ("GO"), doing nothing otherwise ("NO GO"). These last two factors were crossed to produce four task conditions:

ix. positive probe correspondence/positive task
x. positive probe correspondence/negative task
xi. negative probe correspondence/positive task
xii. negative probe correspondence/negative task.

3.2 Dependent variables

Two dependent variables were extracted from the participants' answers:

xiii. Reaction time (in ms)
xiv. Error probability (expressed as the odds, or chance, to observe an answer).

3.3 Operational hypotheses

We hypothesized that reaction time, as well as error probability, would not vary according to target metaphoricity (Hypothesis 1) and target complexity (Hypothesis 2), according to the observations put forth in the introduction.

3.4 Experimental material

The experimental material was designed in Norwegian and contained 80 short sentences of up to 8 words. These sentences were syntactically simple, easy and straightforward to understand. They started with a contextual part and ended with the target word, which carried a figurative or a literal meaning depending on the contextual part. In case of a figurative meaning sentence, the metaphor was carried by the last word of the sentence, for example *Hun tenker* ("She thinks") and *Hun løper* ("She runs"). The sentences did not contain enough information to allow the subject to guess the last word, although they could lead him/her to preselect a semantic domain and a syntactic category for this last word. In order to decide whether the sentence carried a figurative or a literal meaning, the last word had to be read and processed in its context.

40 target words, 3 to 10 characters long, were selected (for example *raskt*, for "quickly") and paired in length and frequency between the two levels of Independent Variable II ("target complexity"). Each target word related to bodily experience (e.g. warm, sweet, heavy, clean, sweet, sour). Each target word was made to complete two paired sentences, giving a literal meaning in one case and a figurative meaning in the other, for example *Hun løper raskt* ("She runs fast," literal meaning) and *Hun tenker raskt* ("She thinks fast," figurative meaning). Half of the target words (= 20) in the figurative condition generated a primary conceptual metaphor, with e.g., *Hun tenker raskt* corresponding to the primary conceptual metaphor INTELLIGENCE IS SPEED. The other half of the target words (= 20) generated a complex conceptual metaphor, such as *Idretten har mange grener* (back-translation: "Sport has many branches"), corresponding to the complex conceptual metaphor SPORT IS A TREE + DISCIPLINES ARE BRANCHES. When designing the material, we strived to associate the context sentence with the target word in such a way that the literal meaning or the figurative meaning would be understood as straightforwardly and easily as possible; for example, we avoided all semantic ambiguity. In addition, we excluded both idioms and novel metaphorical expressions from the material. This last point was particularly important in view of creating homogeneity between primary and complex metaphor conditions, since novel primary metaphors are difficult, if not impossible to collect because of the embodied and universal nature of this category of metaphors. Furthermore, the lack of a standardized method for distinguishing between novel metaphors and conventional metaphors would have required that the level of conventionality be determined for each metaphorical sentence through pre-tests of familiarity conducted among a large number of native speakers. Designing and carrying out such

pre-tests is a whole research programme in itself and beyond the scope of the present research assignment. It must also be mentioned that because the present experiment was synchronized with a neurocognitive experiment (Lachaud and Rydning 2010) that simultaneously measured brain activity, the level of experimental complexity was already considerably high.

The experimental material was pre-tested for the ease and naturalness of comprehension by three Norwegian native speakers unaware of our research goals.

Finally, 80 probing words (words used to probe a given cognitive processing in order to get a measurement), 5 to 13 characters long, were selected and paired in average length and average frequency across conditions. 40 of these probes were related to the meaning of the target word (*raskt*) when presented in a figurative context (e.g. the probe INTELLIGENS for intelligence), and 40 of them were related to the meaning of the target word when presented in a literal context (e.g. the probe HURTIGHET for speed). The same probes were used with figurative and literal items in order to create the two levels of Independent Variable III ("probe correspondence"). Therefore, the meaning of the probe words needed to be unambiguously accepted as related to the figurative meaning or to the literal meaning; they also needed to be unambiguously different from the literal meaning or the figurative meaning, respectively, of the target words to which they were associated in the experiment. The effectiveness of the probe words was pre-tested with Norwegian speakers.

3.5 Experimental plan

4 lists of 40 trials each were created. Each list contained 10 trials of each experimental condition. Half of the trials (5 in each experimental condition) were probed with a positive probe correspondence, half with a negative probe correspondence. Probe words which were used in List 1 were also used in List 4, and probe words used in List 2 were also used in List 3. Each one of the 40 target words was processed by the same individual twice, once in its figurative meaning, once in its literal meaning (i.e. a participant was tested with 80 different sentences). Lists 1 and 2 were used with one and the same participant, while lists 3 and 4 were used with another one, each participant being therefore tested with 80 different probe words. Finally, half of the participants were asked to detect positive probe correspondences, the other half were asked to detect negative probe correspondences.

3.6 Simulation procedure

All the components of a trial, i.e. (i) context sentence, (ii) target word and (iii) probe word, were presented visually at a screen resolution of 800 by 600 and normal pixel density (96 DPI) on a high quality 21" cathode ray tube screen (DELL P1130 Trinitron). The screen was refreshed at 100 Hz, the stimulus delivery being synchronized with screen refreshment for high timing precision of the stimulating procedure. An

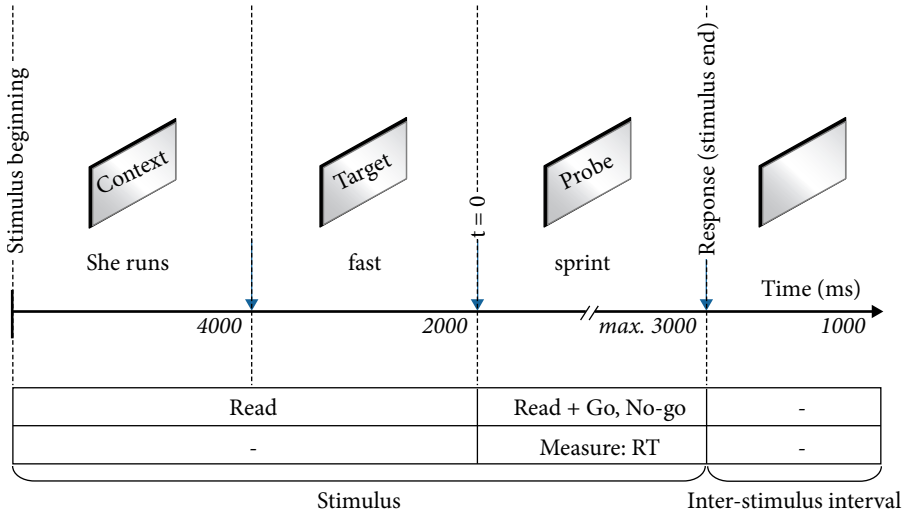


Figure 1. Summary of the stimulation sequence

E-prime script presented trials, randomly selected in a list, until all 40 trials had been processed. Each participant was tested on two lists. The incomplete sentence was displayed for 4 seconds, followed by a fixation cross displayed at the centre of the screen for half a second, followed by the target word presented in isolation at the centre of the screen for 2 seconds. After a blank screen lasting 1 second, the probing word was finally displayed for 3 seconds. The inter-stimulus interval was 1 second. Every 20 trials, the script would enter in pause mode to allow the participant to rest as long as s/he wanted. The stimulation sequence is summarized in Figure 1.

3.7 Answers

Participants could either answer correctly by pressing (or not pressing) the button when supposed to, or make a mistake by pressing (or not pressing) the button when not supposed to, according to the summary of the logical situation given in Table 1 (example extracted from the experimental condition "Figurative/Primary").

3.8 Recording

Reaction times (RT) and responses were recorded using the software E-Prime 2 (Psychological Software Tools, Inc.) and the accompanying serial button box.

Table 1. Answering possibilities of the participants

Task	Meaning	Answer	Error	Context	Target	Probe
Positive	Same	Go	No			INTELLIGENS
		No-go	Yes			
	Different	Go	Yes			SPRINT
		No-go	No	Hun tenker	*raskt*	
Negative	Same	Go	Yes	She thinks	*fast*	INTELLIGENS
		No-go	No			
	Different	Go	No			SPRINT
		No-go	Yes			

3.9 Population

50 young adults (average age 26, min. 20, max. 43, S.D. 5) participated in the study. All were Norwegian native speakers with a good level of English, as assessed by the electronic Dialang test (average score 929, min. 680, max. 1000, S.D. 78). Selecting individuals with a good command of English allowed us to test a population with a language shifting possibility comparable to that of professional translators, in order to extrapolate our results from our population to the population of translators.

3.10 Experimental session and task

The experimental sessions were individual and took place at the Laboratory of the National Centre for Epilepsy at Rikshospitalet, University Hospital, Oslo, Norway. The participant was installed and instructed about the semantic validation task to be performed: read the incomplete sentence, read its completing word, understand the meaning of the completed sentence, read the following word, decide if its meaning is related to the sentence just read, and either press the button of the button box if the answer is yes or no (depending on the participant's group) or do nothing otherwise. The participant was then asked to perform the task as fast as possible and if possible try to avoid making mistakes. Before starting to practise on a set of training items, the participant was asked to summarize what s/he had understood and remembered, and if needed, clarifications were given by the experimenter. The participant was then trained until his or her performance reached a minimum threshold of average speed (1200 ms) and accuracy (80 per cent). The testing part which followed was brief (fifteen minutes maximum) because of its high attentional demand.

3.11 Data processing and analysis

Reaction times and answering odds (i.e. the probability to get a specific answer regarding the association stimulus-task) were processed separately for the positive and negative task conditions. Reaction times for correct answers were filtered with median plus or minus 3 median absolute deviations, by item, and for each condition, by participant. In the positive task condition, 145 values were initially missing due to wrong answers, out of 920 measurements (15.76%). Filtering removed 112 outliers (12.17%). Consequently, the analysis used 663 values (72.07% of the measurements). In the negative task condition, 272 values were initially missing due to wrong answers, out of 1080 measurements (25.18%). Filtering removed 144 outliers (13.33%). Consequently, the analysis used 664 values (61.48% of the measurements).[1] Because answering odds had no missing values, no filtering procedure was necessary to prepare the raw data for the analyses. Answering odds were obtained by simply recording the subjects' behaviour as 0s (wrong answer or outlier) and 1s (correct answer).

Statistical analyses were done using the general linear model, with the software MLwiN 2.1 (Centre for Multilevel Modelling, University of Bristol, UK). The statistical model had 2 random factors (items as Level 1 and subjects as Level 2) in a cross-level structure, and 3 fixed factors (target metaphoricity, target complexity, and the interaction between these two factors) as the experimental variables. Significance was evaluated with large sample chi-squared tests, comparing predicted distributions with and without the variable of interest. Binary data were analysed with a binomial distribution model using the same structure as described above, in order to compute the odds to observe a correct answer depending on experimental variables.

4. Results

The analysis of reaction times neither revealed main effects of target metaphoricity and target complexity, nor interaction between the two factors, whether reaction times were obtained with the positive task or with the negative task (all p-values between .15 and .87).[2] All contrasts between conditions were non-significant (all p-values between .12 and .96).

---

1. The smaller amount of answers in the negative task is due to the higher difficulty of the task, compared to the positive task.

2. A p-value of .15 means that there is a 15% chance of making a mistake when rejecting the null hypothesis. The scientific standard is to accept this risk between 0 and 5%, to consider it fair between 5 and 10%, and to avoid concluding if the risk is over 10%. However, this protective standard has been debated by those who believe that it would be nice to play a lottery where you win 9 times out of 10.

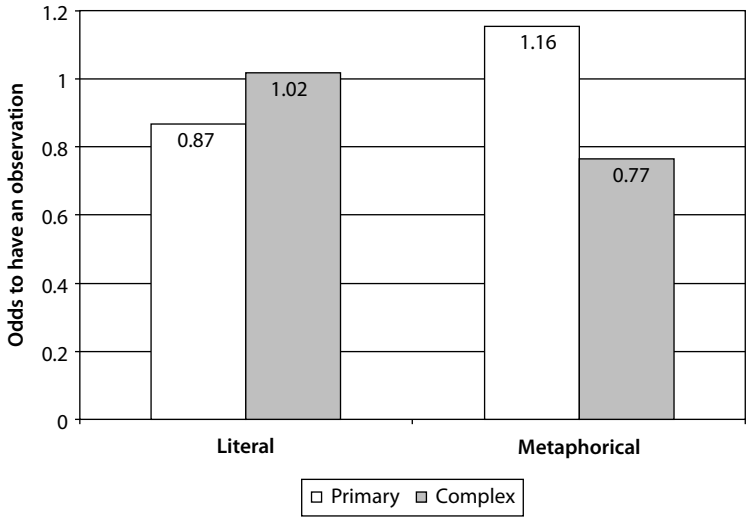


Figure 2. Odds to have an observation per condition for GO task condition to similar meaning

Odds for answering to corresponding meanings in the positive task did not vary significantly depending on either target metaphoricity ($p < .91$) or target complexity ($p < .42$).[3] However, a tendency was observed in the interaction[4] between target metaphoricity and target complexity ($p = .066$). This interaction was mainly due to higher answering odds with primary conceptual metaphors (1.16) than with complex conceptual metaphors (0.77, contrast showing a tendency at $p = .062$ (see Figure 2).

Odds for answering in the reversed task (answering when the meaning is different) varied significantly with target metaphoricity (χ^2 (1, N = 1080) = 13.4; $p < .0003$),[5] but not with target complexity ($p < .11$). Odds to accept different meanings were therefore reduced by 0.46 with a figurative meaning compared to a literal meaning (see Figure 3).

Additionally, a tendency to interaction was observed between target metaphoricity and target complexity ($p < .086$). A further investigation of contrasts reveals that the literal/primary condition was significantly different from the three other conditions

3. These contrasts are respectively between the average of the literal condition ([primary + complex]/2) and the average of the figurative condition ([primary + complex]/2), and between the average of the primary condition ([literal + figurative]/2) and the complex condition ([literal + figurative]/2).

4. Interaction: the variations measured in the dependent variable due to one of the independent variables also depending on the other independent variables.

5. The χ^2 test is the test statistics. In this article, its value (13.4) is only provided for significant contrasts. 1 is the degree of freedom in the independent variable, and N is the sample size.

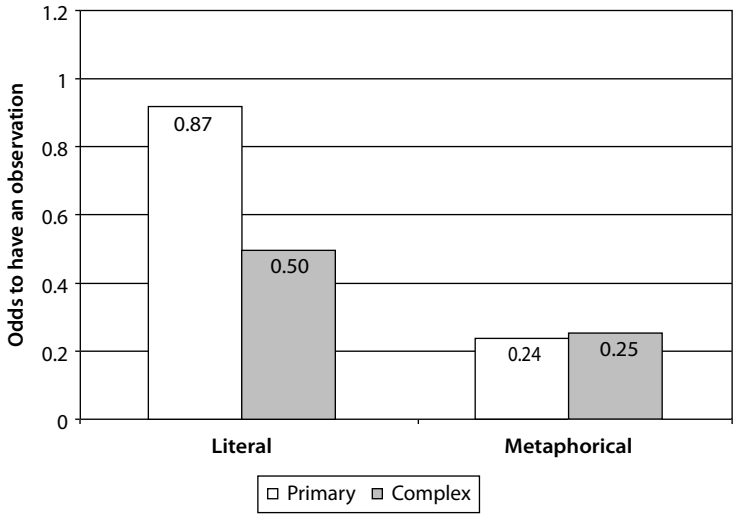


Figure 3. Odds to have an observation per condition for GO task condition to different meaning

(all p-values between .0002 and .02), while these three conditions did not differ significantly from each other (all p-values between .14 and .93).

5. Discussion

The absence of significant reaction time variations depending on target metaphoricity and target complexity was hypothesized (Hypotheses 1 and 2). Once accessed, the meaning of the sentence is probably stored in the person's short-term memory during the probe word's comprehension, before being recalled in the person's working memory, where it is compared with the meaning of the probe word.[6] According to our results, the comparison process took an equal amount of time for detecting meaning similarities or meaning differences, in spite of target metaphoricity and target complexity variations. This seems at first to imply that conceptual clarity is as high in figurative meaning situations as it is in literal situations, and is as high with complex metaphors as it is with primary metaphors.

However, the behavioural pattern in answering odds shows quite a different picture, revealing that conceptual clarity varied depending on target metaphoricity and target complexity, contradicting both our hypotheses and the analysis of reaction times. After double-checking possible technical issues during the processing and

6. For a clarification of the difference between long-term memory, short-term memory and working memory, see Cowan (2008).

analysis of reaction times, such as an adverse effect of distribution filtering, we are confident that the absence of effects in reaction times is not caused by a technical bias. Therefore, the reason for the discrepancy between reaction times and answering odds is to be found somewhere else. Although using the "GO, NO GO" task for its simplicity, robustness, sensitivity, alleged adequacy for this type of study, and large popularity in psycholinguistics, it seems that it did not correctly capture the chronometric dimension of the mental processes we were targeting. This suggestion is reinforced by the fact that some p-values were close to the significance threshold, indicating that some variations depending on the independent variables may have occurred after all, but were not strong enough to be significant.

Therefore, the analysis of answering odds depending on target metaphoricity and target complexity complements our chronometric measurement by revealing interesting cognitive facts which could not have been disclosed by the chronometric investigation of the participants' behaviour alone. If the task was direct (pushing the button when the meaning of the probe word is similar to the meaning of the sentence), there was a greater chance of obtaining a correct answer with primary metaphors than with complex metaphors, indicating a greater certitude of the participant in his or her judgement. No difference was observed with the literal conditions, indicating that the participants judged both situations with even certitude. When the task was reversed, in which case it was more difficult to make a decision (since compared with accepting, rejecting requires additional mental procedures of verification), only an effect of target metaphoricity appeared, with the chance of obtaining an answer being smaller with figurative meaning than with literal meaning – thus showing a greater certitude with the literal meaning than with the figurative meaning. Furthermore, due to the interaction between the two independent variables, it was easier to make a decision with literal primary materials than with materials of other conditions. The conceptual clarity would then be greater in this condition than in the others, including "literal complex." Although a difference between the two literal conditions first appears counterintuitive (theoretically, nothing justifies a distinction between "primary" and "complex" in the literal condition, both being literal), it is interesting to notice that semantic differences may exist, after all. Indeed, it may be the case that the lexical items used for building primary metaphors referred to simpler concepts than the words used for building complex metaphors. However, this would not be an experimental bias due to the selection process of the words during the elaboration of the experimental materials, but would result from the difference in nature of the two types of metaphors studied here, with complex conceptual metaphors possibly only occurring on the basis of more complex concepts than primary metaphors.

Consequently, it seems that conceptual clarity was greater in the mind of our participants (i) with primary metaphors than with complex metaphors, (ii) with literal meaning than with figurative meaning, and (iii) with literal primary meanings than with literal complex meanings. Such results appear to be counterintuitive, and they contradict our hypotheses. However, because of the high statistical reliability of our data (high amount of items and participants, analysis procedure followed), it

is unlikely that our results are incorrect. Therefore, the subjective feeling on which we generated our working hypotheses may not be the reality of the subconscious cognitive processes which were probed and measured in this experiment. Indeed, if there are different types of configurations in language as defined by linguists (literal meaning, figurative meaning in primary conceptual metaphors, figurative meaning in complex conceptual metaphors, and even their literal counterparts), why should the comprehension and production of these specific logical configurations be identical? According to cognitive linguistics, there is a gradient of complexity in the underlying semantic organization, from literal meaning (lower) to complex conceptual metaphor (higher), with primary conceptual metaphors somewhere in between. The behavioural pattern that we observe in our data exactly follows this gradient. Additionally, our synchronized neurolinguistic experiment on brain activity during metaphor comprehension (Lachaud and Rydning 2010) corroborates this finding, showing that brain activity differs significantly between primary and complex conceptual metaphors.

The consequences for translation, of differences as regards how meaning is perceived during comprehension due to target metaphoricity and target complexity, are challenging. If the translator has not properly conceptualized the idea evoked by the linguistic expression during comprehension, his or her approximate conceptual understanding will resurface in the target text, in the form of a non-idiomatic translation. The translator will most probably reproduce the salient features of the original structure, i.e. translate through the device of semantic correspondence. Such an approach, called the transcoding strategy (Seleskovitch 1975), most often turns out to have negative impacts on the outcome of the translation on two levels: it often results in (i) a deviation from the meaning expressed in the original text, leading to obvious shortcomings and translation errors, and (ii) a lack of idiomaticity and immediate intelligibility in the target language, called translationese (Toury 1980), i.e. not necessarily an incorrect translation, but odd, unnatural language. The inherent semantic and structural differences between the two languages only partly explain why it is not always possible to achieve correspondence across languages. More importantly, even though semantic and structural correspondence is achievable in many cases, it is not necessarily the best solution to adopt in a translation, as it may lead to the loss of the spirit of the intended message. Because individuals from different languages and cultures do not always shape their conceptualizations of experience in the same way, salient aspects of a conceptual metaphor focused upon in the source language – used to designate a concept, a person or an object – will not always coincide with the salient aspects of the conceptual metaphor focused upon in the target language and used to designate the same concept, person or object. These preferences, which are the overt results of an individual's ability to construct a conceptual configuration, have direct consequences on translation. The stage of fully understanding the concept evoked by the linguistic formulation in the original text is a prerequisite for the subsequent steps of the translation process, where the translator has to decide on a reformulation strategy.

Conceptual clarity, indeed, could hardly have been objectivized from introspection, from a simple observation or from a detailed examination of the translator's productions and measurable behaviour during production. Therefore, this study brings crucial elements of reflexion for future experimental investigations, refocusing research in Translation Studies on a neglected stage of translation: comprehension and the impact of comprehension on the subsequent steps of the translation process. It is indeed highly challenging to investigate how a translator manages comprehension difficulties during the succeeding translation steps. Additionally, translation mirrors comprehension through the reverbalizing of comprehension. Thus, by testing the assumption that whereas understanding novel metaphors involves meaning creation, understanding conventional metaphors involves meaning retrieval, one could activate and compare in the same experiment these two epistemological factors, which are liable to remain hidden in monolingual experiments.

References

Cowan, N. 2008. "What are the differences between long-term, short-term, and working memory?" In *The Essence of Memory,* W. Sossin, J.-C. Lacaille, V. F. Castelluci and S. Belleville (eds), 323–338. Amsterdam: Elsevier.

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process.* Amsterdam/Philadelphia: John Benjamins.

Fauconnier, G. and Turner, M. 2002. *The Way We Think. Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Gibbs, R. W. 1990. "The process of understanding literary metaphor." *Journal of Literary Semantics* 19: 65–94.

———. 1994. *The Poetics of Mind: Figurative Thought, Language, and Understanding.* Cambridge: Cambridge University Press.

Glucksberg, S. 2003. "The psycholinguistics of metaphors." *Trends in Cognitive Science* 7 (2): 92–96.

Grady, J. 1997. "Foundations of meaning: Primary metaphors and primary scenes." (Doctoral dissertation, University of California, Berkley.)

Grice, H. O. 1975. "Logic and conversation." In *Syntax and Semantics*, P. Cole and J. Morgan (eds), 3: 41–58. New York: Academic Press.

Johnson, C. 1997. "Metaphor vs. conflation in the acquisition of polysemy: The case of SEE." In *Cultural, Typological and Psychological Issues in Cognitive Linguistics,* M. K. Hiraga, C. Sinha and S. Wilcox (eds), 155–169. *Current Issues in Linguistic Theory*. Amsterdam/Philadelphia: John Benjamins.

Kövecses, Z. 2005. *Metaphor and Culture*. New York: Cambridge University Press.

Lachaud, C. and Rydning, A. 2010. "Primary vs. complex metaphors – EEG coherence analysis reveals brain activity differences during metaphor comprehension." Paper delivered at the *Fourth International Conference on Cognitive Science*, June 21–26, 2010. Tomsk, Russia.

Lakoff, G. 1993. "The contemporary theory of metaphor." In *Metaphor and Thought* (second ed.), A. Ortony (ed.), 202–251. Cambridge: Cambridge University Press.

Lakoff, G. and Johnson, M. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
———. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
Lakoff, G. and Turner, M. 1989. *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: University of Chicago.
Liu, B. 2010. Personal communication.
Martikainen, K. 2007. "What can translation reveal of metaphors? Translation experiment on the psychological reality of the conceptual metaphor theory." In *Text, Processes, and Corpora: Research Inspired by Sonja Tirkkonen-Condit*, R. Jääskeläinen, T. Puurtinen and H. P. Stotesbury (eds), *Publications of the Savonlinna School of Translation Studies* 5, 145–159. Joensuu: Joensuu University Press.
Narayanan, S. 1997. "Embodiment in language understanding: Sensory-motor representations for metaphoric reasoning about event descriptions." (Doctoral dissertation, Department of Computer Science, University of California, Berkley.)
Searle, J. 1979. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
Seleskovitch, D. 1975. *Language, langues et mémoire. Étude de la prise de notes en interprétation consécutive*. Paris: Lettres Modernes Minard.
Toury, G. 1980. *In Search of a Theory of Translation*. Tel Aviv: The Porter Institute for Poetics and Semiotics.

Innovative subtitling

A reception study*

Alexander Künzli and Maureen Ehrensberger-Dow
Zurich University of Applied Sciences, Switzerland

This paper presents the results of an experimental study investigating reception capacity and audience response to subtitled movies. Twenty-seven viewers were shown four movie excerpts, with commercially available standard subtitling or with innovative subtitling. The latter comprised additional information regarding language and culture-specific elements in the original soundtrack. Data were collected simultaneously with eye-tracking and consecutively with questionnaires. Eye-tracking measured fixation duration and percentage of gaze time in the various areas of interest, whereas the questionnaire assessed accuracy on questions about movie content and audience perception and satisfaction. The results show no significant differences in accuracy between the two conditions, indicating that viewers of subtitled audiovisual productions are able to process more information than established subtitling norms suggest.

Keywords: subtitling, reception capacity, audience response, eye-tracking, questionnaire

1. Introduction

Dialectal and sociolectal elements, ethnolect, and slang are a few examples of linguistic varieties that characterize oral discourse. They can have a large number of functions in feature as well as in documentary films, from conveying local color or the background of a particular character to illustrating social and hierarchical power relationships. Independent of the question as to the extent that the language in any film actually reflects spontaneous orality, certain features of original film dialogues

* We would like to thank Peter Jud for helping with the preparation of the materials, Beate Sick for providing advice on the experimental design and statistics, the IUED students for participating in this study, and two anonymous reviewers for their valuable comments and suggestions. Any errors remain our own.

are generally considered difficult or impossible to render in subtitles because of space and time constraints. Many studies on subtitling (e.g. Hurt, Koloszar and Lisa 1997; Reinart 2004; Georgakopoulou 2009) have therefore identified a leveling of speech varieties such as dialects or sociolects. Such leveling is even explicitly called for in handbooks for subtitlers. Ivarsson and Carroll (1998: 75), for example, argue that subtitlers should make subtitles as inconspicuous as possible so as not to distract the audience from the film itself – a norm most subtitlers seem to follow. A similar tendency of striving for inconspicuousness has often been observed regarding the handling of extralinguistic culture-bound elements (see Nedergaard-Larsen 1993 for a definition of intra- and extralinguistic culture-bound translation problems and, more recently, Pedersen 2007), which are often left untranslated, probably in an effort to avoid any negative feedback effect from the original soundtrack.

However, little is actually known about the cognitive effort necessary to watch subtitled movies. Even the legitimacy of the often cited norm, according to which a subtitle of two full lines with 80 characters should be displayed for about 6 seconds (see Ivarsson and Carroll 1998: 67), still awaits empirical testing. D'Ydewalle and Gielen (1992: 416), for instance, rightly argue that such a presentation time may be preferred by audiences simply because they are used to reading subtitles of that duration. The acceptance of the norm of inconspicuous subtitling is also being challenged by more innovative, experimental forms of audiovisual translation such as "fansubs" (i.e. amateur subtitling on the internet). Fansubs violate mainstream subtitling norms in several respects: the use of subtitles of more than two lines, notes at the top of the screen and glosses in the body of the subtitles to explain cultural referents, and capitalization of entire words to represent intonation are just a few examples. The main aim of these alternative practices is to allow a more comprehensive understanding of the original film dialogues (Díaz Cintas and Muñoz Sánchez 2006: 45).

The need for more research into reception is well recognized in Translation Studies and audiovisual translation. Gambier (2008: 30) recently criticized the fact that in subtitling research, reception studies almost exclusively focus on intralingual subtitling for the deaf and hard-of-hearing. One exception is Caffrey (2008), who has reported results from a study on interlingual subtitling, investigating the cognitive effort necessary to watch Japanese TV animes subtitled in English with either standard subtitling or with additional pop-up glosses. The use of pop-up gloss was suggested to have increased processing effort, as indicated by various measures from the eye-tracking data, such as a larger proportion of skipped subtitles and a significantly lower gaze time in the subtitle area. These objective measures were corroborated by the subjective assessments given by the participants in an electronic questionnaire, indicating that they experienced the same subtitles as faster when pop-up glosses were used. However, it is not clear whether it was the additional information on the screen or the mode of presentation of that information that resulted in the apparent increase in cognitive effort.

The empirical study reported in this paper therefore sets out to investigate reception of and audience response to a different type of innovative subtitling more

similar to standard subtitling. Responses to movie extracts with standard, commercially available subtitling were compared by means of a mixed method design to the extracts with additional information provided in "surtitles" at the top of the screen. Indicators of the cognitive effort necessary to watch subtitled movie excerpts were collected simultaneously with eye-tracking technology and consecutively by means of a questionnaire. The questionnaire was designed to provide information on the participants' subjective perception of subtitle speed, their satisfaction with innovative subtitling, and their retention of different types of visual and verbal information in the movie excerpts to be viewed. The study is part of a larger research project being carried out to investigate the production and reception processes of subtitled movies.

By measuring whether visual and verbal information in movies is better retained with the established presentation norms of subtitling than with more innovative practices of subtitling, and by assessing the acceptance of such innovative practices, our study addresses the central issue of the effect of reading subtitles on audience satisfaction and the allocation of audience attention. The focus here is thus not on the impact of innovative subtitles on more complex psychological and cognitive processes of overall movie comprehension (see Gambier 2003: 185 for different research strands into subtitle reception). Rather, the focus is on attitudinal and perceptual issues, i.e. the viewers' preferences for standard or innovative subtitling, and on their subtitle-decoding strategies.

2. Method

2.1 Participants

Twenty-seven undergraduate students (23 women and 4 men, mean age = 27.7 years) enrolled in a translation program at Zurich University of Applied Sciences volunteered to participate in the present experiment. They were contacted by e-mail and informed that the study aimed at investigating the reception of subtitled movies. All participants were native German speakers. The participants were randomly assigned to one of two groups (A or B; see Section 2.3 below), except for the restriction that there be an equal number of men in each group.

Before viewing the film, the participants were asked about their competence in the language of the movie to be viewed (French). On a 5-point scale (1 = "very good" and 5 = "weak"), the average self-reported judgment of listening comprehension of French was 2.5, with no significant difference between the two groups. Only two participants reported having already seen the film (one in each group). Moreover, one may assume that the participants are regularly exposed to subtitles – in larger Swiss cities there is generally a choice of cinemas offering either subtitled or dubbed versions – but not necessarily to innovative subtitles such as fansubs.

2.2 Material

Four excerpts between 59 and 138 seconds long were chosen from the French movie *Entre les murs* ("The Class," 2008, directed by Laurent Cantet), which is commercially available with German subtitles. The movie is a semi-autobiographical account of a middle school teacher of French language and literature in a suburb of Paris. The pupils are aged from 13 to 15 and are of mixed ethnic and cultural background. The movie and two excerpts of it were chosen because the dialogues contain a high degree of linguistic varieties in the form of sociolect, ethnolect, slang or *verlan*, which is typically used by young French speakers and consists of inverting syllables in a word. The two other excerpts were chosen because they contain potential extralinguistic culture-bound translation problems (i.e. references to France's geography, history, or media). An analysis of the German subtitles had shown that the intralinguistic culture-bound elements were generally either leveled or omitted altogether, whereas the extralinguistic culture-bound references were rendered verbatim, without any potentially useful explicitations for the target audience.

Two versions of each excerpt were therefore prepared, one with standard subtitling (French soundtrack and commercially available German subtitles) and the other with additional information in German in the upper part of the screen (i.e. surtitles[1]), made for the occasion to supplement the commercially available subtitled version. The surtitles contained metalinguistic information such as explanations of word play, of connotations of linguistic items that were rendered in the German subtitles only in their denotative sense, if at all, and of culturally specific elements. The design resulted in two conditions for each of the four excerpts, one containing surtitles (the treatment or "sur" condition), and one without surtitles (the normal or "sub" condition). An example of screen captures from both conditions can be seen in Figure 1: the surtitle in the screenshot on the right explains the unintended pun produced by one of the pupils, which the teacher interprets as an impudent comment. Surtitles in Excerpt 2 provided additional information such as details about French history and on the French philosopher Voltaire's novels *Candide* and *Zadig*. Surtitles in Excerpt 3 dealt mainly with the connotations of the French verb form *imparfait du subjonctif*. Excerpt 4, finally, contained among other things additional information on French TV series mentioned in the original soundtrack (*Tropiques amers* and *Sécurité intérieure*). The selection of the segments of the original soundtrack to be surtitled and the wording of the surtitles were done in teamwork by a translator and a subtitler with more than 10 years of professional experience.

Three measures of text density were determined for each excerpt (see Table 1). The first was the number of surtitles and subtitles: each excerpt had 5 or 6 surtitles, and between 12 and 22 subtitles. The second measure of text density was the number of characters per subtitle or surtitle: on average, the length of the surtitles (Ø = 103.0

1. The term "surtitle" in this paper is thus not used in its traditional sense (i.e. for the opera and the theater).

(a) subtitle condition (b) surtitle condition

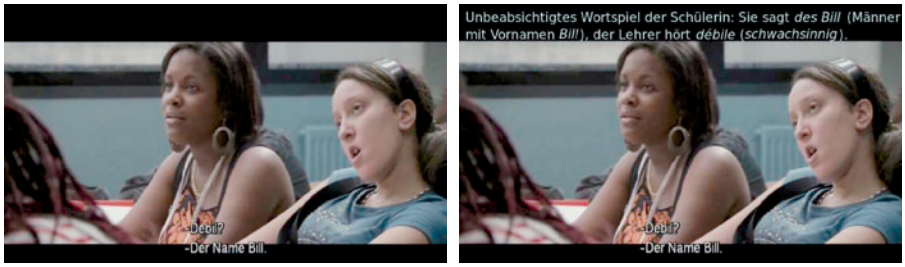


Figure 1. Example of the third sequence of Excerpt 1 in the two presentation conditions

characters) was higher than the conventional recommendation of a maximum of 80 characters for subtitles and more than three times as long as the average length of the commercially available subtitles in the movie excerpts used in the present study (Ø = 33.3 characters). The third measure of text density was the number of characters per minute to be read by the participants in each of the four clips. The average total of 1802.5 characters per minute in the sur condition was more than twice as much as the reading speed of 750 characters per minute applied for instance by the European Captioning Institute (Georgakopoulou 2009: 34).

2.3 Procedure

Data collection took place in the usability laboratory at Zurich University of Applied Sciences. All of the participants in the study were tested individually and under similar conditions. Upon arrival at the laboratory, they were asked to read and sign an

Table 1. Text density in each excerpt (in terms of number of subtitles, surtitles and characters)

	Excerpt 1	**Excerpt 2**	**Excerpt 3**	**Excerpt 4**	
Number of:					Total
Subtitles	22	16	13	12	63
Surtitles	6	5	5	5	21
Total (sur+sub)	28	21	18	17	84
Average number of characters in:					Overall
Subtitles	33.9	29.3	33.2	37.4	33.3
Surtitles	95.7	98.0	125.4	94.6	103.0
Number of characters per minute in:					Average
Subtitles	846.1	833.0	711.7	858.0	812.2
Surtitles	846.5	801.5	1245.0	1068.3	990.3

Table 2. Presentation conditions (italics is *sub* condition, bold is **sur** condition)

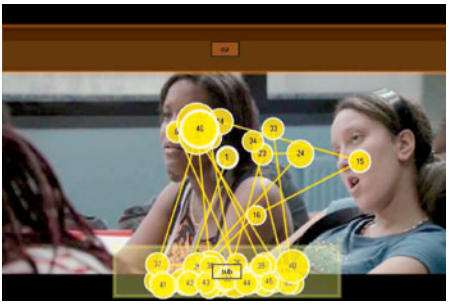
	Excerpt 1	Excerpt 2	Excerpt 3	Excerpt 4
Group A (n = 14)	*sub*	**sur**	*sub*	**sur**
Group B (n = 13)	**sur**	*sub*	**sur**	*sub*

Note: *sub* = normal condition (subtitles only); **sur** = treatment condition (surtitles in addition to subtitles).

informed consent form. They were then seated in front of an eye-tracking monitor. After the calibration procedure, the participants answered some preliminary questions on-screen (e.g. age, sex, level of French competence) and then were asked to watch the four movie excerpts. Each participant saw the excerpts in alternating "sub" and "sur" conditions, as outlined in Table 2, depending on which group they were assigned to.

While each excerpt was viewed, various types of simultaneous data were collected by the Tobii TX60 eye-tracker and computed by the Tobii Studio software.[2] The eye-tracking equipment consisted of a flat-panel computer monitor with special, discrete diodes mounted around the edges. By tracking the movements of the eye, they record what part of the screen a person is attending to at any particular time. The eye-tracking data available include: number and duration of fixations over 30 ms,[3] gaze time in various areas of interest (i.e. subtitle, surtitle and the rest of the screen image), delay until the first fixation, shifts between areas of interest, skipped sub- and surtitles, probability of word fixation, pupil size, etc. For the purposes of the present discussion, only the first three measures will be examined in detail (see Figure 2 for examples from two participants of the areas of interest, recorded fixations, and gaze paths for the third sequence of Excerpt 1, which lasted approximately 12 seconds).

(a) subtitle condition (participant a22)



(b) surtitle condition (participant b21)

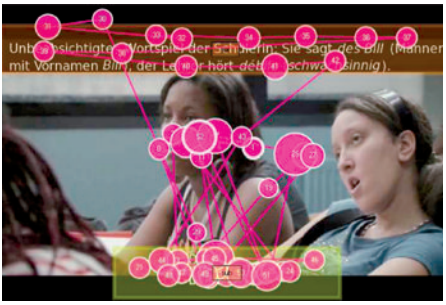


Figure 2. Fixations and gaze paths for the third sequence of Excerpt 1 in the two presentation conditions

2. Tobii Studio version 1.7.2; see http://www.tobii.com/scientific_research.aspx.

3. This minimum fixation duration was chosen to allow comparisons with Caffrey (2008).

After each excerpt, consecutive data was collected by means of on-screen questions about the participants' reading and viewing strategies (i.e. measures of accuracy) and their perception of and reaction to the sub- and surtitles (i.e. judgments of speed and usefulness). Specifically, there were five questions for each excerpt that assessed reception capacity with regard to three types of information (cf. Hickethier 2007): two about the images (i.e. iconic attention), two about the dialogue (i.e. verbal attention), and one about the overall situation in the excerpt (narrative attention). For each of these questions, there were five answer options: one correct choice, three incorrect choices, and a final option of "I don't know." The questions were presented in the same order to all participants. In addition, two questions addressed participants' perceptions of the subtitle and surtitle speed as well as the usefulness of the additional metalinguistic information given in the surtitles (if present). After the final excerpt, participants were asked to provide their overall judgments about the surtitles.

3. Results

The results concerning reception capacity as well as audience response to the subtitles and surtitles are presented first and then compared with the on-line measures provided by the eye-tracking data.

3.1 Reception capacity

The participants were able to answer 68.3% of the questions about the movie content correctly. The order of presentation of the conditions did not seem to matter, since the groups performed very similarly, although there was a slight difference in the proportion of "don't know" answers (11.1% for group A and 8.1% for group B). Most strikingly, however, there were no differences in overall accuracy based on presentation condition (sub = 67.8%; sur = 68.8%) or on excerpt (see Table 3). The additional surtitles therefore did not seem to have had a distracting effect on the participants' attention processes and reception capacity.

The analysis of the question types revealed differences, however, with answers about the general situation in the excerpt most accurate and questions about the image least accurate (see Table 4). Although questions about the dialogue were slightly

Table 3. Percentage correct by group, excerpt, and presentation condition (italics is *sub* condition, bold is **sur** condition)

	Excerpt 1	**Excerpt 2**	**Excerpt 3**	**Excerpt 4**	**Overall**
Group A	*64.3*	**70.0**	*70.0*	**70.0**	68.6
Group B	**70.8**	*64.6*	**64.6**	*72.3*	68.1
Overall	67.5	67.3	67.3	71.2	68.3

Table 4. Percentage correct by question type and presentation condition (italics is *sub* condition, bold is **sur** condition)

Question type:	***sub***	**sur**	**Overall**
Image	*50.6*	**55.6**	53.1
Dialogue	*74.3*	**70.3**	72.3
Situation	*89.1*	**92.4**	90.8

Table 5. Subjective rating of time available by presentation condition (italics is *sub* condition, bold is **sur** condition) and usefulness of surtitles

	Excerpt 1	**Excerpt 2**	**Excerpt 3**	**Excerpt 4**	**Overall**
Time available*:					
sub	*4.2*	*4.4*	*4.6*	*4.9*	*4.5*
sur	**3.5**	**3.4**	**2.8**	**3.9**	**3.4**
Overall	3.9	3.9	3.7	4.4	4.0
Helpfulness** of:					
surtitles	**3.3**	**3.7**	**2.9**	**3.5**	**3.4**

* 1 = not enough time, 5 = more than enough time.
** 1 = very helpful and 5 = very distracting.

more accurate in the subtitle only condition and those about the image and situation in general slightly more accurate in the condition including surtitles, none of these differences were significant according to Fischer's exact test for count data.

3.2 Audience perception and satisfaction of subtitles and surtitles

The judgments made after each excerpt of the time available to read the text on the screen were quite positive overall for both conditions (4.5 and 3.4 for the sub and sur condition, respectively, on a scale from 1 = not enough time, 5 = more than enough time), although the judgments for the condition including surtitles were significantly lower than those for the subtitles only condition ($p < 0.001$, Fischer's exact test for count data; see Table 5). This pattern held for all of the excerpts, although only the difference for Excerpt 3 was highly significant.

The qualitative judgments of the helpfulness of the metalinguistic information provided in each of the excerpts for the surtitle condition were rather neutral, with an overall average for the four excerpts of 3.4 (where 1 = very helpful and 5 = very distracting). There was little difference in judgments between the excerpts (see Table 5). A possible explanation for this result could be the fact that the participants had on the whole quite good knowledge of French (see Section 2.1) and were therefore less dependent on the surtitles.

For the question at the end of the experiment about the usefulness of surtitles in general, the average judgment across participants of 3.3 (where 1 = very useful, 5 = very distracting) was very similar to their judgments of the metalinguistic information provided in the surtitles, with a suggestive but non-significant difference between the two groups (overall judgments for Group A = 3.6 and for Group B = 3.0). Interestingly, there was no correlation between judgments of the usefulness of surtitles in each excerpt and of the time available to read them (rank correlation = 0.046).

3.3 Eye-tracking results

The first two on-line measures of attention (mean fixation length and number of fixations) revealed a slight difference between the two presentations. The mean fixation length for each participant for each excerpt was calculated by dividing the sum of all fixations over 30 ms by the number of fixations. Since some of the eye-tracking records suggested that certain participants spent very little time fixated on the screen for certain excerpts, these data points were eliminated. Specifying that a minimum of 50% of the total time had to be fixated on the screen meant that all data from two participants (one per group) and data for certain excerpts for other participants had to be eliminated (two additional excerpts from the sub condition and four from the sur condition). Using this criterion, there was a significant difference in mean fixation length between the two conditions (sub = 0.36 s; sur = 0.32 s, $p < 0.01$), although the difference in the actual number of fixations between the two conditions did not reach significance (sub = 227; sur = 251; see Table 6). There were also no significant differences in the mean fixation length for the four excerpts. The differences in the number of fixations for each extract were predictable, since they were directly related to the length of each excerpt. The t-tests used for these comparisons were for independent samples with unequal variance.

The third measure of attention from the eye-tracking data was gaze time in various areas of interest. Three areas of interest were defined: (1) the area at the bottom

Table 6. Mean fixation length (meanFlength in seconds) and number of fixations (allFcount) by presentation condition and excerpt (italics is *sub* condition, bold is **sur** condition)

	Excerpt	Excerpt 1	Excerpt 2	Excerpt 3	Excerpt 4	Overall
sub	*meanFlength*	*0.34*	*0.35*	*0.37*	*0.38*	*0.36*
	allFcount	*274*	*345*	*134*	*165*	*227*
sur	**meanFlength**	**0.32**	**0.31**	**0.32**	**0.33**	**0.32**
	allFcount	**283**	**373**	**156**	**197**	**251**
Overall	meanFlength	0.33	0.33	0.35	0.36	0.34
	allFcount	279	358	145	180	238
Length		108 s	138 s	59 s	73 s	

Table 7. Percentage of gaze time in different areas of interest by presentation condition (italics is *sub* condition, bold is **sur** condition)

	Subtitle area	Surtitle area	Rest of image
sub	*31.5%*	*0.1%*	*68.5%*
sur	**29.4%**	**10.8%**	**59.8%**
Overall	30.4%	5.4%	64.2%

of the screen where the subtitles were shown; (2) the area at the top of the screen where the surtitles were shown in the surtitle condition; and (3) the rest of the screen image (see Figure 2 for examples). Again, a criterion of at least 50% of overall time spent looking at the screen was set, which meant that two excerpts for one participant had to be eliminated (one in each condition). An examination of the proportion of gaze time in the three areas of interest showed that all except one participant looked at subtitles in all the excerpts (and she did in three of the four excerpts). All except one (other) participant must have noticed the surtitles as well, because they looked at them for some portion of the time. The analysis indicated that significantly less time was spent gazing at the image if surtitles were present ($p < 0.01$, t-test for independent samples with unequal variance; see Table 7).

Whereas the amount of time spent reading the subtitles was similar in the two conditions, participants spent significantly less time looking at the visual nonverbal image when additional metalinguistic explanations were displayed. This result has to be weighed against the fact that retention of visual elements was, however, better in the surtitle condition, as Table 4 revealed.

4. Discussion

The aim of this study was to investigate reception of and audience response to subtitled movies by testing the legitimacy of the commonly accepted norm in the subtitling industry and literature, according to which a subtitle should not exceed two lines of text and a maximum of 80 characters. If two full lines are used, the subtitle should be present on screen for six seconds. Shorter subtitles are displayed accordingly (see, for example, Ivarsson and Carroll 1998). Moreover, it is assumed in the subtitling industry that average reading speed should not exceed 750 characters per minute (Georgakopoulou 2009).

In a mixed method design, reading and viewing strategies of university students who were asked to watch several movie excerpts were studied in two different conditions. A questionnaire tested accuracy with respect to three types of information: iconic attention (the image), verbal attention (the dialogue), and narrative attention (the general situation). Eye-tracking, on the other hand, measured percentage of gaze time in different screen areas as well as mean fixation durations. Subtitling was either according to the existing norm (i.e. commercially available standard subtitling, with

one or two lines of subtitled text) or with additional metalinguistic explanations (i.e. innovative subtitling that included both subtitles and surtitles). The latter is already practiced in amateur subtitling on the internet. The surtitles explained the use and connotation of culturally specific references in the original French soundtrack and of elements that belonged to non-standard linguistic varieties and that were either leveled or not translated in the commercially available German subtitled version of the movie.

The results of the study challenge the legitimacy of the above-mentioned subtitling norm. Watching a movie without subtitles involves an almost continuous distribution of attention between two input channels and two types of information, visual and auditory. Subtitles add a third source of information to this already complex situation. In the present study, a fourth information source was added. The presence of surtitles, comprising metalinguistic explanations, did not negatively affect the participants' reception capacity. This result suggests that movie watchers are able to process a larger amount of textual information without being distracted from the plot than is generally assumed.

There are some important implications from our results. First, viewers, and perhaps particularly younger cohorts of viewers, might be able to process more information in the form of subtitles and surtitles than has been conventionally assumed. If more information can be included in text form, then a deeper understanding of the original movie might be possible. And finally, surtitles would be an important step in tailoring such information to the needs of specific audiences. For example, the deaf and hard-of-hearing, language teachers, or learners might be prepared to pay for such additional information in the form of surtitles.

The established rules regarding presentation time and maximum number of characters per subtitle seem however to have been internalized by the participants in the present study. Thus, when asked to assess the time available to read the text in standard vs. innovative subtitling, participants' ratings differed significantly between the two modes. The amount of time available for processing text in the condition with surtitles was experienced as significantly lower than in the condition with subtitles only, suggesting an additional cognitive load. However, participants' performance in terms of retention of various verbal and visual elements in the movie excerpts was identical in the two conditions. This result suggests that standard subtitle presentation rules may simply be preferred because viewers are used to seeing subtitles at that duration and length.

The study also involved three types of eye-movement measures: number of fixations, mean fixation duration per excerpt, and percentage of gaze time in the various areas of interest (i.e. subtitles, surtitles, and rest of image). The results reveal a significant difference in mean fixation duration between excerpts with subtitles only and those with both subtitles and surtitles. Fixations were shorter for excerpts that had surtitles but there was no significant difference between the average number of fixations in the two conditions, suggesting that the participants spent more time moving between various areas of interest when an excerpt included surtitles and less time

fixating on individual points. This may be related to a higher cognitive load due to the presence of additional text on the screen, which is also reflected in the more negative judgments of the time available to process the textual information.

The analysis of eye-tracking data by areas of interest indicated a significant difference regarding the percentage of gaze time in the areas of surtitles, image and subtitles. Surprisingly, especially in light of the results reported by Caffrey (2008), the time spent gazing at surtitles in the present study was not at the cost of time spent on the subtitles, because less time was spent gazing at the rest of the image in the condition with surtitles. This is also unexpected in light of the questionnaire data, which indicated differences in accuracy based on question type (image, dialogue, general situation), but no significant differences between the two conditions. In the surtitle condition, accuracy was slightly lower for questions regarding the retention of elements that concerned the dialogue, yet actually slightly higher for the questions regarding the general situation and the image. The last result seems counterintuitive: it suggests that although less time was spent looking at the image in the surtitle condition, visual information was retained better in that condition than in the subtitle condition. Although not statistically significant and therefore possibly attributable to chance, the difference between the two conditions might be due to attention-switching between the bottom and top of the screen when surtitles are present and the processing of additional visual information as eyes sweep vertically instead of merely in the horizontal direction of text. Introducing textual information in two different areas of the screen might therefore actually improve overall processing of images. This hypothesis needs further empirical investigation, however.

We are aware of the possible limitations related to our findings. First, the participants in our study are university students and therefore represent a fairly young and literate cohort. It can be assumed that they are able to handle a higher reading speed than other demographic groups. On the other hand, at least in Switzerland, moviegoers have a choice between dubbed and subtitled versions. Those who decide to watch a movie with subtitles may have higher reading skills in general, and additional metalinguistic information might not distract them from enjoying a movie. Moreover, they represent a generation that has virtually grown up with audiovisual media. Established subtitling norms may not meet their needs, expectations, and capacities in every respect. Secondly, the movie excerpts prepared for the experiment were fairly short. High reading speeds might be unproblematic in the short run, whereas watching a whole movie with additional surtitles could lead to considerable fatigue and thus reduced reception capacity. This possible limitation has however to be weighed against the fact that the clips represented segments of the original soundtrack that contained a large amount of language- and culture-specific references. Not all scenes of an entire movie would need surtitles, of course. Finally, it has to be acknowledged that there was generally little on-screen activity going on while the surtitles were displayed. The polysemiotics of audiovisual productions is a variable to be taken into consideration in empirical research into screen translation. In movies where the visual nonverbal components play a decisive part, additional text might distract viewers from the plot.

On the other hand, one may claim that surtitles are particularly relevant for art films ("writers' films") such as the one used in this experiment. Writers' films tend to favor the verbal message over the other sign systems that make up the audiovisual text (Zabalbeascoa 2008).

Audiovisual media play a major role in the modern world, at home, at school, and at work. It is therefore only natural that research into audiovisual translation is a rapidly expanding domain. So far, however, only a few studies have addressed the issue of reception, as Gambier (2008) recently pointed out. The aim of the present study was to make a contribution in this direction, by investigating how people are able to divide and shift their attention in such a complex situation as watching subtitled movies. Moreover, by testing the reception and acceptance of subtitles that violate established norms, we hope to have contributed to helping define new research paths exploring viewers' actual needs and capacities as well as the readability and legibility of subtitled audiovisual productions. The present study focused on the handling of culturally specific elements and of non-standard language use, a fundamental challenge since subtitling involves the shift between two modes that are each governed by their own rules – the oral and the written – under space and time constraints. Other challenges await further research, in particular the question of how surtitles containing additional metalinguistic information should best be formulated and presented so as to boost movie comprehension. The segmentation of subtitles, their position on the screen, and the use of color to indicate paraverbal information are other areas that need to be explored further.

References

Caffrey, C. 2008. "Using pupillometric, fixation-based and subjective measures to measure the processing effort experienced when viewing subtitled TV anime with pop-up gloss." In *Looking at Eyes. Eye-tracking Studies of Reading and Translation Processing,* S. Göpferich, A. L. Jakobsen and I. M. Mees (eds), 125–144. Copenhagen: Samfundslitteratur.

Díaz Cintas, J. and Muñoz Sánchez, P. 2006. "Fansubs: Audiovisual translation in an amateur environment." *The Journal of Specialised Translation* 6: 37–52.

D'Ydewalle, G. and Gielen, I. 1992. "Attention allocation with overlapping sound, image, and text." In *Eye Movements and Visual Cognition: Scene Perception and Reading*, K. Rayner (ed.), 415–427. New York: Springer.

Gambier, Y. 2003. "Introduction. Screen transadaptation: Perception and reception." *The Translator* 9 (2): 171–189.

———. 2008. "Recent developments and challenges in audiovisual translation research." In *Between Text and Image. Updating Research in Screen Translation,* D. Chiaro, C. Heiss and C. Bucaria (eds), 11–33. Amsterdam/Philadelphia: John Benjamins.

Georgakopoulou, P. 2009. "Subtitling for the DVD industry." In *Audiovisual Translation: Language Transfer on Screen,* J. Díaz Cintas and G. Anderman (eds), 21–35. New York: Palgrave.

Hickethier, K. 2007. *Film- und Fernsehanalyse.* Stuttgart: Metzler Verlag.

Hurt, C., Koloszar, C. and Lisa, C. 1997. "Transfer zwischen Kunst und Technik – Zur Problematik der multimedialen Übersetzung." In *Text – Kultur – Kommunikation: Translation als Forschungsaufgabe; Festschrift aus Anlass des 50jährigen Bestehens des Instituts für Übersetzer- und Dolmetscherausbildung an der Universität Graz,* N. Grbič and M. Wolf (eds), 167–181. Tübingen: Stauffenburg.

Ivarsson, J. and Carroll, M. 1998. *Subtitling.* Simrishamn: Transedit.

Nedergaard-Larsen, B. 1993. "Culture-bound problems in subtitling." *Perspectives: Studies in Translatology* 1 (2): 207–241.

Pedersen, J. 2007. "Scandinavian subtitles: A comparative study of subtitling norms in Sweden and Denmark with a focus on extralinguistic cultural references." (Doctoral dissertation, Stockholm University, Sweden.)

Reinart, S. 2004. "Zu Theorie und Praxis von Untertitelung und Synchronisation." In *Literarisches und mediales Übersetzen. Aufsätze zu Theorie und Praxis einer gelehrten Kunst,* R. Kohlmayer and W. Pöckl (eds), 73–112. Frankfurt am Main: Peter Lang.

Zabalbeascoa, P. 2008. "The nature of the audiovisual text and its parameters." In *The Didactics of Audiovisual Translation,* J. Díaz Cintas (ed.), 21–37. Amsterdam/Philadelphia: John Benjamins.

Errors, omissions and infelicities in broadcast interpreting

Preliminary findings from a case study

Daniel Gile
ESIT, Université Paris 3 Sorbonne Nouvelle, France

Recordings and transcripts of simultaneous interpretation outputs of President Obama's inaugural speech as broadcast by TV stations in French, German and Japanese were scrutinized and analyzed for errors, omissions and infelicities in an investigation of cognitive saturation and language-pair specific difficulties. Inter alia, it was found that many misinterpreted or omitted source-speech micro-units were simple and non-technical and that there were more errors and omissions in the Japanese renderings than in either the German or French renderings. Findings also suggest that there may be different interpreting styles in terms of preference given to language correctness versus information completeness. All these are consistent with the tightrope hypothesis, according to which interpreters tend to work close to cognitive saturation, which also makes language-specific and language-pair specific idiosyncrasies relevant parameters in the interpreting process. Findings also suggest that simple methodology and naturalistic studies can make valuable contributions in Interpreting Studies.

Keywords: errors, omissions, infelicities, tightrope hypothesis, language-pair specificity, naturalistic research, methodology

1. Introduction – Experimental and sophisticated versus naturalistic and simple

Experimental research is hailed as an advanced form of investigation that makes it possible to target specific phenomena and the effects of specific variables by controlling the relevant local environment. In research into translation, the most widespread experimental paradigms over the past two decades have been think-aloud protocols (TAPs) and computer logging, with gaze analyzers becoming increasingly popular (see for instance Göpferich 2008). In her work on explicitation in translation, Englund

Dimitrova (2005) chose to use triangulation, focusing on TAPs and computer logs. In Chapter 3 of her book, she discusses the validity of TAPs – as does Toury (1995: 234–238) – and highlights doubts about this paradigm in the research community. TAPs cannot be implemented in interpreting, where other experimental online and quasi-online methods have been used, including measurements of blood pressure, heart rate, skin conductance, pupil diameter, EEG patterns, brain activation, cortisol levels in the blood to name just a few (Tommola and Hyöna 1990; Klonowicz 1994; Kurz 1995; European Parliament 2006; Ahrens, Kalderon, Krick and Reith 2010). All these methods raise issues of ecological validity because of the artificial nature of the environment in which phenomena are observed and/or because of their intrusiveness, hence the importance of naturalistic data for cross-checking.

Another aspect of recent empirical research into both translation and interpreting is the increasing use of sophisticated methods (including complex experimental designs) and technology (especially for brain imaging and other physiological measurements). These require considerable know-how, and some require expensive equipment as well. And yet, Translation Studies is mostly conducted within either translator and interpreter training programs or modern language departments as opposed to research centers with significant resources, and many investigators are translation or interpreting instructors with little or no training in research methods or graduate students working with them. It seems to me (as was already the case in the 1980s – see Gile 1990), that for such colleagues, simple methods using naturalistic data can be a good introduction to empirical research and an opportunity to hone their skills before attempting more sophisticated studies. What is perhaps equally important is that such simple methods are still likely to yield an abundance of solid, meaningful findings that can then be tested and fine-tuned with other designs, thus contributing significantly to TS.

With regard to research into interpreting, until recently, few authentic source and target speeches were available for public scrutiny, and those available were generally confined to single interpretations into single languages. Nowadays, important speeches, press conferences and press briefings are regularly broadcast by television stations with simultaneous interpretation, each station with its own interpreters, and both the original delivery and several interpretations can often be found on the internet, including the large database of European Parliament speeches interpreted into all the EU languages (see http://www.europarl.europa.eu/activities/plenary/cre.do?language=EN). This provides opportunities for comparisons. One pioneering comparative study was Strolz's doctoral dissertation on three German target speeches (Strolz 1992) – see also Pöchhacker (1997) and Neuberger (2010). This paper presents the initial findings from an ongoing study. It is an attempt to explore the potential of the analysis of easily available naturalistic data to detect trends and test hypotheses with little technology and simple methods that do not necessarily follow standard formal procedures.

2. The material and the situation

The material is taken from live French, German and Japanese interpreting renditions of US President Barack Obama's inaugural speech delivered on January 20, 2009, on French, German, Swiss, Austrian and Japanese television stations (France 2, LCI, TF1 and Public Sénat for France; Arte, SF1 and ZDF for German renditions; and NHK, CNBC, TBS, Fuji and CNN Japan for Japan).

The text of the original speech is available on many internet sites (e.g. http://www.nytimes.com/2009/01/20/us/politics/20text-obama.html), and video recordings can also be found on the internet (e.g. http://www.youtube.com/watch?v=VjnygQ02aW4). The speech was delivered at a speed of about 130 words per minute on average, though when taking into account several pauses of several seconds, the actual rate during individual speaking chunks could be somewhat higher (2413 words in about 18 minutes and 30 seconds). AIIC recommends, somewhat vaguely, an ideal presentation rate of about 3 minutes per 40-line page of speech (http://www.aiic.net/ViewPage.cfm/article14.htm), and more generally in the literature, a speech delivery rate of about 120 wpm is considered "normal" – for instance, in Shlesinger (2000), it was selected as the lower rate for a comparison with a faster rate of 140 wpm. This would make Obama's inaugural speech somewhat faster than average.

Broadcast interpreting is a professional specialization for interpreters in some countries, in particular in Japan, where it has often been analyzed in the literature (see for example BS Hōsōtsūyakugurūpu 1998). In European countries, it tends to be done by regular conference interpreters. While in some cases, specific norms apply to broadcast interpreting as opposed to conference interpreting, in the case of Obama's speech, besides a high stress level (see Kurz 2002, 2003), I will assume that the work and the interpreters' priorities were similar to those applicable in other assignments where speeches of important personalities are interpreted either on- or off-site. Some parts of the speech were difficult to translate because of information density and style, metaphors, biblical/historical allusions and citations, but many other parts pose no particular difficulty in terms of language or encyclopedic knowledge.

The sample includes five veteran interpreters working into French, five working into German and six working into Japanese. The analysis in this paper looks essentially at the performance of

- three experienced interpreters (who had graduated from highly selective graduate training programs and/or had extensive experience interpreting speeches in English by scientists, intellectuals and Heads of State with more than 15 years of experience) working into their French A with a text (the performance of one younger interpreter working without a text will only be mentioned in passing);
- two interpreters working into German with a text (the third interpreter worked without a text, while the transcripts of the last two interpreters were only received a week before the manuscript of this paper was sent to the editors and could not be included); on the basis of both the audio files (kindly provided by Franz

Pöchhacker), including their voice, prosody and overall performance (few hesitations, false starts, pronunciation problems, etc.), and the circumstances (an important speech interpreted in countries where professional conference interpreters work regularly on similar occasions), I will assume that they were also experienced professional interpreters working into their A-language (this and other parameters should be checked further later in the project);

– five interpreters working into Japanese with a text (working conditions were different for the sixth team, which interpreted the speech after one day of preparation); according to Chikako Tsuruta, a veteran broadcast interpreter and professor at Tokyo University of Foreign Studies, who kindly sent me the Japanese transcripts, all five Japanese interpreters were professionals as well, with Japanese A (native language) and English B (active foreign language) and more than 15 years of experience.

The following important limitations arising from the context apply and have been taken into account when selecting the methods for processing the data:

1. My A-language is French; while it seems plausible that I can reliably spot blatant deviations from linguistic acceptability (grammatical errors, gross mispronunciations, glaring infelicities) in the French versions, no such claim can be made as regards the Japanese and German versions.
2. For Japanese, only the transcripts were available, and I could not check them against the recordings.
3. Working conditions were not identical in all cases. Three of the French interpreters and two of the interpreters working into German had the text 20 to 40 minutes ahead of the actual (live) address. In Japan, all the interpreters are reported to have had the text 20 to 40 minutes before the speech, and worked in teams of two. Chikako Tsuruta reports that one team divided up the speech between the two interpreters during preparation and believes that the same was done in the other teams, as is standard practice among broadcast interpreters in Japan (personal communication). With such an arrangement, in relative terms, preparation time for the Japanese interpreters would be at least double that available for the other interpreters (same time but half the text for the interpreter translating the first part of the speech, more time for the second half of the text for the second interpreter).

3. Errors, omissions and infelicities, Effort Models and the Tightrope Hypothesis

The conceptual framework selected for this study can be described briefly as follows.

3.1 The Effort Models and the Tightrope Hypothesis

The "Effort Model" for simultaneous interpreting (see an up-to-date discussion in Gile 2009) is a conceptual framework in which cognitive operations involved in simultaneous interpreting are pooled into three core "Efforts": the Listening and Analysis Effort L, the Production Effort P (including self-monitoring) and the short-term Memory Effort M. A Coordination Effort C is also required for processing capacity management. Simultaneous Interpreting is represented as follows (the "plus" sign is to be interpreted in a wide sense, not as an arithmetic sum of numerical values):

(1) SI = L + M + P + C

It is assumed that each of the Efforts requires some processing capacity, and that in spite of possible "cooperation" between them, the simultaneous operation of any two Efforts or of the three core Efforts entails an additional cognitive cost above the cognitive cost of any single Effort. Most importantly, it is assumed that total processing capacity requirements during simultaneous interpreting are close to total available capacity (the "Tightrope Hypothesis" – see Gile 1999b, 2009).

In simultaneous interpreting with text (formula 2), which corresponds to the type of interpreting investigated in this study, the additional Reading Effort R is assumed to cooperate with the Listening Effort (and reduce its processing capacity requirements), but it also competes with it insofar as speakers often deviate from their text, and that interpreters need to keep their attention on both the incoming speech sounds and the text and compare the two so as to advance at the right speed and detect possible discrepancies. Moreover, in simultaneous interpreting with text, the constant visual presence of source language words and structures before the interpreter's eyes is assumed to induce higher risks of language interference, and target language production is assumed to require further processing capacity to keep the target speech free of such interference.

(2) SI with text = L + R + M + P + C

For smooth interpreting, at any time, total available processing capacity in the interpreter's brain needs to be larger than or equal to total processing capacity requirements; in addition, processing capacity allotted to each Effort at any time needs to be sufficient to cover the needs associated with the task it is engaged in.

When one or more of these conditions is/are not met, performance in one or more Effort(s) deteriorates, which may give rise to errors, omissions and/or infelicities. EOIs can therefore reflect not only insufficient knowledge of the relevant languages and/or insufficient extralinguistic knowledge, but also saturation due to high processing capacity requirements as well as processing capacity management errors.

While the particular pooling of cognitive operations into "Efforts" chosen for the Model is only a convention, the Tightrope Hypothesis is a "theory" that requires empirical support. Such support can come from several phenomena, such as:

1. Many mistranslations and omissions unlikely to result from insufficient source speech comprehension;
2. An unusually high frequency of infelicities;
3. When comparing several interpreters' renderings of the same source speech, a high spread of EOIs (different units of the source speech are affected in each interpreter's output) – such evidence was found in an experimental study with ten interpreters (Gile 1999b), and again in a replication by Matysiak (2001);
4. When scrutinizing the same interpreter's first and second rendition of the same source speech, the presence of EOIs in the second run that were not present in the first suggested that the lack of linguistic and extralinguistic knowledge to understand and reformulate the relevant segment adequately could not be incriminated; such evidence was also found in Gile (1999b).

3.2 Causal relationships in EOIs

Cognitive saturation can affect any and all of the Efforts, either directly (for instance, when insufficient availability for listening leads to incomprehension of an incoming speech segment), or indirectly, through carry-over effects (see Gile 2009), and it is difficult to attribute individual EOIs to specific causes. However:

1. A strong norm taught in conference interpreting schools is grammatical correctness in the interpreter's output. Professionals tend to prefer omissions to ungrammatical or unfinished sentences. When they feel cognitive load becomes too heavy for the completion of all tasks required to produce a satisfactory output, they may decide to omit part of the information and keep their speech grammatically correct; omissions are often (but not necessarily) the result of tactics, whereas grammatical errors are the result of loss of control. A large number of grammatical errors in an interpreter's output suggests a loss of control more strongly than omissions in an output that suffers from no grammatical errors.
2. Long omissions (long clauses, full sentences and above) are often associated with tactical choices, especially in simultaneous interpreting with text, when the interpreter is lagging behind and may decide that the loss associated with such an omission is less damaging than the risk of lagging even further behind and missing more important information.
3. While some blurring of consonants and vowels is a "natural" secondary effect of high articulation speed, blatant pronunciation errors when articulation speed is moderate can also be interpreted as reflecting loss of control.
4. Meaning errors can result from insufficient background knowledge or linguistic knowledge, or from signal distortions (the speaker's strong unfamiliar accent, background noise), from cognitive saturation affecting the Listening Effort, or, more interestingly, from a processing capacity deficit in the Production Effort. In non-technical everyday language (as opposed to rare words or specialized terms),

false cognates in a skilled interpreter's output into an A language (native language) are likely to result from such cognitive failures.

Clearly, distinctions can be subtle and difficult to document reliably in a given recording or transcript of an interpreter's output. They are given here essentially as an indication of how veteran conference interpreters and interpreter trainers tend to interpret the relevant phenomena in the field and in the classroom. Over many years of experience in examination committees that discuss the performance of interpreting students and interpreters point by point, I have found that such interpretations are highly convergent.

4. Research questions

This study seeks tentative answers to the following questions:

1. In this corpus, are there enough EOIs not likely to result from the interpreters' insufficient linguistic or extralinguistic knowledge to suggest that at least in the interpretation of similar speeches under similar circumstances, they are a regular feature of interpreting performance?
2. Are EOI patterns similar for all interpreters or do they differ by individual? By target language? Is there evidence supporting the idea that interpreting in some language combinations is more difficult than in others?

In addition, this study seeks to assess the power of analyzing the interpreting output at micro-unit level (the procedure is explained later in this paper) to identify patterns and trends.

5. Materials and method

5.1 Speeches and comparisons

- The French versions found on the Internet were transcribed (by myself), and both the recordings and the transcripts were used for analysis. In view of the objectives of this study, as well as my linguistic limitations and the fact that I would depend on others for other transcripts, my transcription was limited to plain text and intonation-based punctuation (commas and periods), with no marks for prosody and only triple-dot marks for what I perceived as hesitation pauses.
- German recordings were compared "online" to a transcript of Obama's speech in order to spot blatant errors and omissions. A comparison of the transcript of the original speech with German transcripts from Neuberger 2010 was done later, when they became available thanks to Franz Pöchhacker. The errors and omissions I identified in both comparisons were the same (see below).

– Only transcripts were used for the analysis of the Japanese speeches, as I did not have access to the recordings.

5.2 Sensitivity level

One important strategic decision was where to set the sensitivity level. High sensitivity detection tools pick up "weak" signals, but often at the cost of a higher risk of mistaking "noise" for "signal," that is, erroneously identifying irrelevant phenomena as significant. Evidence from previous studies suggests that there is high inter-individual variability in sensitivity to and norms of linguistic acceptability and fidelity (Gile 1985, 1995, 1999a). In view of the nature of the questions being investigated here, it seemed reasonable to prioritize reliability and choose a low sensitivity level. I therefore only listed what seemed to me clear meaning errors, language errors (including blatant infelicities) and omissions of significant bits of information. I did not at this stage count all infelicities or omissions in cases where I thought that a particular component of the source text was possibly missing because the interpreter deemed it insignificant, for instance when saying "we" for "America" or omitting "Today" and just using the present tense.

Setting the sensitivity threshold at such a low level means that part of the relevant information may not be picked up in the analysis. It may be taken up at a later stage of the project. On the other hand, the probability that EOIs identified are "true EOIs" is enhanced. A few tests were conducted to check this: in a research methods classroom, a one-page extract of the transcript of one French interpreter's rendering was distributed to nine students who were asked to identify language errors and infelicities (*maladresses de langue* in French). There was some variation in the reports of the four native French speakers in the class, but all four included all my EOIs. In addition, over several weeks, I showed colleague interpreters different pages of transcripts of both the original and target texts in French and German and asked them to identify meaning errors and omissions (*fautes de sens* and *omissions*), or asked them about individual segments and their translations. All their EOIs invariably included mine. For Japanese, I asked (graduate) students of Japanese-French-English interpreting to identify errors and omissions in Japanese renderings on the basis of a table with a transcript of the original speech (see Table 2 for an illustration of the micro-units principle) and of transcripts of the Japanese target-language versions and checked them against mine. Three reports were returned. Again, all the EO's I identified were also identified by the students (no infelicities were involved in this analysis).

This procedure may appear casual, and indeed it would not be sufficient if an attempt was made to identify all EOIs for finer analyses of interpreting quality. As explained above, in view of the objectives of this study, a low sensitivity level was selected and the checking of my own EOI diagnoses was performed essentially as additional insurance against what could have been idiosyncratic norms. By asking a relatively large number of people (essentially colleagues with whom I was working

in the booth) their opinion on a small number of cases each time, I managed to have more confirmations than would have been possible with a planned procedure under strictly controlled conditions.

6. Findings

6.1 In the French sample

Table 1 shows the relative frequency of four categories of errors and infelicities. Language errors are grammatical errors and blatant infelicities; meaning errors are target speech segments where it seemed clear that the interpreter had misunderstood the idea expressed by a word or group of words (see examples below). In Table 1, every omission is counted once whether it affects a single word, a clause or a whole sentence (another analysis is offered later with metrics of the amount of information affected).

In the recordings of the four French interpreters, the distribution of EOIs was found as shown in Table 1.

Here are a few examples of EOIs in the French outputs:

Language errors (and blatant infelicities)

1. When translating "Homes have been lost, jobs shed": *Des foyers ont été perdus, des emplois disparu* (a verb is missing between *emplois* and *disparu*). The faulty French grammatical structure parallels the English structure, which suggests a lack of processing capacity for Production.
2. When translating "our schools fail too many": *Trop des enfants échouent à l'école* (*des enfants* should be *d'enfants*).
3. When translating "this is the price and the promise": *Tel est le prix et la promesse*. This is a grammatical agreement problem, and the sentence should read: *Tels sont le prix et la promesse*…. The same grammatical lack of agreement is found in the source speech, which suggests a lack of processing capacity for monitoring and correction in Production as in Example 1.

Table 1. Distribution of EOIs in the French-A sample

Type of EOI / interpreter	F1	F2	F3	F4	F4'
Language error	7	14	5	22	32
Meaning error	6	11	5	26	37
Omission	12	34	48	34	49
Phonological distortion	2	5	3	7	10
Total	27	64	61	89	128

Note: Only the first 13 minutes of the speech are available for F4. The numbers in F4' are a rough linear extrapolation of what they could be if he had interpreted the whole speech. They are obviously speculative and only indicated to give readers an idea of the magnitude of differences between the performance of the 4 interpreters.

4. When translating "the father of our nation ordered these words to be read to the people": *A annoncé que les mots suivants soient lus au peuple. A annoncé* means "announced," and does not "agree" with the subjunctive in *soient lus*. The sentence would make sense and be grammatically correct if *a annoncé* were to be replaced with *a ordonné* (ordered). The interpreter may have thought of *a ordonné* while pronouncing *a annoncé*. Again, this seems to indicate insufficient processing capacity for monitoring and correction in Production, as in Examples 1 and 3.

Omissions

1. "far-reaching" (… network)
2. "enduring" (… spirit)
3. "hard" (in "plowed the hard earth")
4. "of standing pat" (in "the time of standing pat")
5. "vital" ("… trust between a people and the government")
6. "new" ("… threats")
7. "and understanding" (in "greater cooperation and understanding between nations")
8. "hard-earned" ("… peace in Afghanistan")

Meaning errors

1. *Pour eux* ("for them," whereas the speaker says "for us"). An insufficient knowledge of English can reasonably be ruled out as the source of this error.
2. *Nous produirons des merveilles technologiques* ("we will produce technological wonders," for "and will wield technology's wonders" ["… to raise healthcare's quality and lower its cost"]. In this case, it is unlikely, but not impossible that this highly experienced interpreter had problems understanding the use of "wield."
3. *PNB* ("GNP" for "GDP").

Salient phonological error

1. *Responsibilité* (for *responsabilité*). The incorrect vowel corresponds to the English vowel in "responsibility," suggesting processing capacity saturation and the inability to inhibit interference from the source language.
2. *Le foyage* (for *le voyage*). There is no obvious explanation for the phonological error here.

On the basis of these figures, the following comments can be made:

Roughly, in a speech of less than 19 minutes, there were somewhat more than 1 EOI per minute in F1, and more than three EOIs/min in F2 and F3. In other words, blatant errors, omissions and infelicities were not exceptional occurrences in a speech made at a relatively high delivery speed, but with the interpreters having had some time to study the text beforehand. (F4's EOIs are even more numerous, but since he was far less experienced than the other interpreters and had no text, his performance is not taken into account in this comment).

Table 2. An illustrative extract from the general table identifying errors and omissions (marked 1) on micro-units in Obama's speech by the four French interpreters

Source speech segment / interpreter	F1	F2	F3	F4
Rather, it has been the risk-takers,				
the doers, the makers of things,				
some celebrated				
but more often				
men and women				
obscure in their labor,				
who have carried us			1	
up the long		1	1	
rugged	1		1	
path			1	
towards prosperity			1	
and freedom.			1	1
For us	1			
they packed up				
their few			1	
worldly possessions			1	
and travelled across oceans				
in search of a new life.				
For us				
they toiled in sweatshops			1	
and settled the West,		1		1
endured the lash of the whip				
and plowed				
the hard	1			1
earth.				

F1, F2 and F3, who had graduated from highly selective graduate training programs and/or had extensive experience interpreting speeches in English by scientists, intellectuals and Heads of State, may reasonably be assumed to have excellent comprehension of English. They may conceivably not be too familiar with some words and idioms in Obama's speech such as "when the levees break," may have failed to identify the reference to Paul's Letter to the Corinthians in "the time has come to set aside childish things" or wondered about the meaning of "country" in "the city and the country," but the vast majority of errors and omissions in their performance affected simple words and grammatical structures (see examples in Table 2 and in Section 6.3).

EOI performances vary considerably from one interpreter to the next. At this low level of sensitivity, their number goes from 27 (for F1) to a maximum estimated at four times that number (for F4), with F2 and F3 in-between. EOI "styles" also seem

to vary from one interpreter to the next. F2 makes many language and meaning errors (about 40% of the total number of EOs, vs. 53% of omissions), while F3 makes far fewer errors (about 16% of the total number of EOs, vs. 79% of omissions). This suggests that F3's behavior is more tactical in that he prefers to omit segments rather than risk making language and meaning errors, while F2 tries to translate everything and makes more numerous errors. This impression is strengthened by the fact that infelicities are numerous in F2 and rarer in F3 – but not being at this point in a position to show reliable evidence of that difference, I will keep this comment anecdotal. Note that in F4, language errors, meaning errors and omissions are more evenly distributed, with a high number of language errors, which are probably symptomatic of a loss of control.

Having compared several renderings with the original and noted that errors and omissions often affected single words and in particular adjectives, I then divided the source speech into micro-units at clause and sub-clause level for the purpose of more detailed analysis. Table 2 is an extract from the overall table drawn up and illustrates the data. For every micro-unit, the presence (marked 1) of an error or omission (infelicities and language errors were not taken into account) was checked for each interpreter's rendering. By adding the 1s in each column, the total number of micro-units affected could be calculated. For instance, in the extract in Table 2, 3 micro-units were affected in F1, 2 units were affected in F2, 9 units were affected in F3 and 3 units were affected in F4. This procedure allows one to calculate the relative informational weight of each omission or error (in F3, there were only 3 instances of omissions, but one affected 6 micro-units and one affected 2 micro-units). Moreover, the tables indicate how many interpreters stumbled on each micro-unit, thus helping probe deeper into the phenomenon (see Section 6.3).

When this table for the first five minutes of the speech, up to and including "For us, they fought and died, in places like Concord and Gettysburg, Normandy and Khe Sahn," is divided into 212 micro-units, the results for the four French interpreters (Table 3) show a different pattern from that found in Table 1.

Interestingly, if the total number of EOIs were to be taken as a quality indicator, Table 1 suggests a ranking with F1 as best, followed in turn by F3, F2 and F4; if the total number of micro-units of the source speech affected by EOs is the criterion, F1 is still the best for the first five minutes of Obama's speech, followed by F2, F4 (the one recent graduate who had no text) and F3 (now fourth instead of second, essentially because his omissions are long and tend to affect several micro-units at a time instead of one).

Table 3. A comparison between the total number of EOs and the number of affected micro-units in the first five minutes of Obama's inaugural speech – the case of the French interpreters

Interpreter	**F1**	**F2**	**F3**	**F4**
EOs in the whole speech for F1, F2, F3, in the first 13 minutes for F4	18	45	53	60
Number of affected micro-units in the first 5 minutes	5	10	26	18

Also note (Table 1) that F3 makes the smallest number of language errors, and his recording shows a small number of phonological distortions, while F2 makes more and F4 many more. To listeners, F3 might well have been the most pleasant interpreter to listen to after F1, as may be checked in a later phase of the project.

6.2 Inter-language comparisons

The issue of language-specificity in interpreting is controversial. Some theorists claim that if interpreters have the required skills, language-specific and language-pair specific factors such as differing syntactic structures should have little or no impact on interpreting difficulty. Setton (1999) reports that in his trilingual corpus with English, German and Chinese, he found no obvious effects of syntactic differences in interpreting between different language pairs. Another view is that because of high processing capacity requirements in interpreting, such differences are significant (see a more detailed discussion of the issue in Gile 2009).

Syntactically, lexically and culturally, French could be considered closest to English, followed by German and then by Japanese (see Wilss 1978; Kurz 1983; Ito-Bergerot 2006; Kondo 2008). Two aspects of EOIs will be considered here: the overall number of micro-units affected by EOs in each target-language sample, and the number of "long omissions" (affecting several successive micro-units of the source speech).

Table 4 shows the number of micro-units in the first 5 minutes of source speech affected in each rendition. F1 to F3 are French interpreters and their renditions, G1 and G2 are German-speaking interpreters and their renditions, J1 to J5 are Japanese-speaking interpreters and their renditions:

The mean number of affected micro-units in these first 5 minutes of the speech is lowest in French (with high inter-individual variability), and almost doubles on average in the German version. Scores are also highly variable in the Japanese sample, with J1 and J5 within the range of values for the French sample, but J2, J3 and J4 far above them. In view of the assumptions, mentioned earlier, that the Japanese interpreters were all professionals with at least 15 years' experience and that they had more time to prepare for their part of the speech than the French and German interpreters, these data seem to support the idea that language-pair-specific differences can indeed have an impact on the difficulty of interpreting.

Table 4. The number of micro-units of the first five minutes of Obama's speech affected by EOs in French, German and Japanese renditions

Interpreter	**F1**	**F2**	**F3**	**G1**	**G2**	**J1**	**J2**	**J3**	**J4**	**J5**
Number of micro-units affected by EOs	5	10	26	25	26	13	73	50	57	18
Mean value by language		13.7			25.5			42.2		

Table 5. Number of long omissions by interpreter and by language

Interpreter	F1	F2	F3	G1	G2	J1	J2	J3	J4	J5
Number of chunks ≥ 3 micro-units	0	1	4	0	3	0	8	3	8	0
Mean by language		1.7			1.5			3.8		
Number of chunks ≥ 4 micro-units	0	0	3	0	2	0	4	3	4	0
Mean by language		1.0			1.0			2.2		
Number of chunks ≥ 5 micro-units	0	0	2	0	0	0	4	2	3	0
Mean by language		0.7			0			1.8		

Analysis of EOs by the number of long omissions ("chunks" of 3, 4 and 5 consecutive micro-units respectively) generates a clear "hierarchy" between French and German on one hand and Japanese on the other, but not between French and German (Table 5). Since the samples are very small and because of other uncontrolled variability factors, these can only be considered tentative indications, especially in view of the high inter-individual variability in each target-language sample, but they do suggest that simultaneous interpreting from English into Japanese may be cognitively more taxing than work into German or French.

6.3 Frequent EOs

Finally, in connection with the question of the relative "interpreting difficulty" of speech segments affected by EOIs: in the first five minutes of the speech, 132 micro-units were omitted or mistranslated by the 10 experienced interpreters who worked with the text into their native language. As can be seen from Table 6, more than a third of them are found in no more than one interpreter's rendition, about two thirds (66%) in no more than two interpreters, and 91% in no more than four interpreters, which indicates a large spread of the EOs over the speech and replicates one finding of Gile (1999b).

As to the micro-units most affected by EOs, they are:

1. "fail too many" (in "our schools fail too many"), in eight interpreters out of ten
2. "long" (in "who have carried us up the long, rugged path"), in seven interpreters
3. "far-reaching" (in "a far-reaching network"), in six interpreters
4. "the ways we use" (in "the ways we use energy"), in six interpreters
5. "rugged" (in "who have carried us up the long, rugged path"), in six interpreters
6. "hard" (in "plowed the hard earth"), in six interpreters.

As can be seen, none of these micro-units are linguistically difficult or require special thematic knowledge, though understanding exactly what is meant by some of them may require some analysis, and finding an appropriate equivalent in the target language may also cost effort.

Table 6. Number of interpreters who omitted or mistranslated each micro-unit in the first 5 minutes of the speech

Number of EOs found in … interpreter(s) out of 10	**1 int.**	**2 ints.**	**3 ints.**	**4 ints.**	**5 ints.**	**6 ints.**	**7 ints.**	**8 ints.**	**> 8 ints.**
Number of EOs	50	37	21	12	7	4	1	1	0
Percentage of EOs	38%	28%	16%	9%	5%	3%	.8%	.8%	0%

As indicated in Section 3.1, all these findings are consistent with the idea that processing capacity management difficulties and/or cognitive saturation, as opposed to poor linguistic and/extralinguistic knowledge, cause many EOs.

7. Methodological comments and conclusion

The research procedure followed here leaves some uncertainty: Ideally, all relevant details about each interpreter and about the specific conditions under which s/he prepared for the assignment and interpreted the speech, including the physical environment, available help from a colleague on site, type of headset used, type of monitor available to view President Obama's speech, quality of the sound, etc., would have been known and ascertained as comparable. Ideally, all transcripts would have been performed under and checked against a set of unique transcription rules by a team of native speakers of French, German and Japanese.

Also, as mentioned earlier, the low sensitivity level chosen for EOI detection here can only detect part of the picture, and subtle differences in the behavior of both individual interpreters and language groups may well have gone unnoticed.

Nevertheless, in view of the research questions addressed, I believe that above the noise left by uncertainty, the data that emerge allow the following tentative conclusions:

1. The large number of EOIs that were not likely to be due to insufficient linguistic or extralinguistic knowledge is consistent with the hypothesis that cognitive saturation is involved.
2. Different interpreting "styles" with respect to priority given to completeness of information versus high linguistic quality of the output may be characteristic of specific interpreters.
3. Differences in the numbers of EOs and long omissions suggest that it was more difficult to interpret the English speech into Japanese than into French or German (differences between French and German renderings need to be investigated further). This lends some empirical support to the idea that under the heavy cognitive pressure of interpreting, linguistic idiosyncrasies and language-pair specific idiosyncrasies do matter.

Going beyond this specific study towards research policy:

1. For practical reasons, it would be difficult to organize an experiment in which the output of French, German-speaking and Japanese interpreters could be examined as was done in this study. If only for this reason, the simple, naturalistic method chosen here, which can easily be applied to other languages and provide interesting data that would be difficult to collect otherwise, deserves attention.
2. The type of investigation described here can be replicated easily by individual researchers with virtually no technical and financial means other than a computer with access to the internet. The large amount of material available on the internet in many languages, as well as individual recordings of TV interpretations, open up possibilities for many case-study type replications that can be more powerful in answering questions than a small number of experiments.
3. Finally, performing this type of study for graduation theses can be appealing to interpreting students who tend to shy away from research, because it involves detailed scrutiny of genuine interpretations that they can use for learning purposes and never takes them into highly abstract theoretical considerations. My own interpreting students were happy to look at the transcripts and help, though at ESIT they are not required to conduct research.

References

Ahrens, B., Kalderon, E., Krick, C. M. and Reith, W. 2010. "fMRI for exploring simultaneous interpreting." In *Why Translation Studies Matters*, D. Gile, G. Hansen and N. K. Pokorn (eds), 237–248. Amsterdam/Philadelphia: John Benjamins.

AIIC. *Guideline for Speakers*. http://www.aiic.net/ViewPage.cfm/article14.htm. (Accessed June 12, 2010).

BS Hōsōtsūyakugurūpu (BS 放送通訳グループ). 1998. *Hōsōtsūyaku no sekai* (放送通訳の世界) (*The World of Broadcast Interpreting*). Tokyo: Ark.

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process*. Amsterdam/Philadelphia: John Benjamins.

European Parliament. 2006. *Report on the 3rd Remote Interpretation Test, Brussels, 22.11–1.12.2005*. Executive Summary.

Gile, D. 1985. "La sensibilité aux écarts de langue et la sélection d'informateurs dans l'analyse d'erreurs : une expérience." *The Incorporated Linguist* 24 (1): 29–32.

———. 1990. "Research proposals for interpreters." In *Aspects of Applied and Experimental Research on Conference Interpretation*, L. Gran and C. Taylor (eds), 226–236. Udine: Campanotto Editore.

———. 1995. "Fidelity assessment in consecutive interpretation: an experiment." *Target* 7 (1): 151–164.

———. 1999a. "Variation in the perception of fidelity in simultaneous interpreting." *Hermes* 22: 153–172.

———. 1999b. "Testing the effort model's tightrope hypothesis in simultaneous interpreting – a contribution." *Hermes* 23: 51–79.

———. 2009. *Basic Concepts and Models for Interpreter and Translator Training*. Revised Edition. Amsterdam/Philadelphia: John Benjamins.

Göpferich, S. 2008. *Translationsprozessforschung. Stand – Methoden – Perspektiven*. Tübingen: Gunter Narr Verlag.

Ito-Bergerot, H. 2006. "Le processus cognitif de la compréhension en interprétation consécutive : acquisition des compétences chez les étudiants de la section japonaise." (Doctoral dissertation, ESIT, Université de la Sorbonne Nouvelle, France.)

Klonowicz, T. 1994. "Putting one's heart into simultaneous interpretation." In *Bridging the Gap. Empirical Research in Simultaneous Interpretation*, S. Lambert and B. Moser-Mercer (eds), 213–224. Amsterdam/Philadelphia: John Benjamins.

Kondo, M. 2008. "Multiple layers of meaning – towards a deepening of the "sense" theory of interpreting." In *Towards Quality Interpretation in the 21st Century. Proceedings of the 6th National Conference and International Forum on Interpreting*, E. Wang and D. Z. Wang (eds), 35–40. Beijing: Foreign Language Teaching and Research Press.

Kurz, I. 1983. "'Der von uns…': Schwierigkeiten des Simultandolmetschens Deutsch-Englisch." In *Festschrift zum 40 Jährigen Bestehen des Instituts für Übersetzer- und Dolmetscherausbildung der Universität Wien*, 91–98. Tulln: Ott Verlag.

———. 1995. "Watching the brain at work – an exploratory study of EEG changes during simultaneous interpreting (SI)." *The Interpreter's Newsletter* 6: 3–16.

———. 2002. "Stressfulness of live TV interpreting vs. conference interpreting." In *Translation: New Ideas for a New Century. Proceedings of the XXVI FIT Congress, Vancouver, BC, Canada, August 7–10, 2002*. 202–206.

———. 2003. "Live TV interpreting. A high-wire act?" In *Avances en la investigación sobre interpretación*, Á. Collados Aís and J. A. Sabio Pinilla (eds), 159–171. Granada: Comares.

Matysiak, A. 2001. "Controlled processing in simultaneous interpretation: A study based on Daniel Gile's Effort Models." (Master's thesis, University of Poznan, Poland.)

Neuberger, B. 2010. "Mediendolmetschen und Rhetorik – eine korpusbasierte Analyse." (Master's thesis, University of Vienna, Austria.)

Pöchhacker, F. 1997. "Clinton speaks German: A case study of live broadcast simultaneous interpreting." In *Translation as Intercultural Communication – Selected Papers from the EST Congress – Prague 1995*, M. Snell-Hornby, Z. Jettmarova and K. Kaindl (eds), 207–216. Amsterdam/Philadelphia: John Benjamins.

Setton, R. 1999. *Simultaneous Interpretation*: *A Cognitive-Pragmatic Analysis.* Amsterdam/Philadelphia: John Benjamins.

Shlesinger, M. 2000. "Strategic allocation of working memory and other attentional resources in simultaneous interpreting." (Doctoral dissertation, Bar Ilan University, Israel.)

Strolz, B. 1992. "Theorie und Praxis des Simultandolmetschens. Argumente für einen kontextuellen top-down Ansatz der Verarbeitung und Produktion von Sprache." (Doctoral dissertation, Vienna University, Austria.)

Tommola, J. and Hyönä, J. 1990. "Mental load in listening, speech shadowing and simultaneous interpreting: A pupillometric study." In *Foreign Language Comprehension and Production*, J. Tommola (ed.), 179–188. AFinLA Yearbook, Turku Publication of the Finnish Association for Applied Linguistics 48.

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam/Philadelphia: John Benjamins.

Wilss, W. 1978. "Syntactic anticipation in German-English simultaneous interpreting." In *Language Interpretation and Communication*, D. Gerver and H. Wallace Sinaiko (eds), 343–352. New York/London: Plenum Press.

On cognitive processes during wordplay translation

Students translating adversarial humor

Ulf Norberg
Stockholm University, Sweden

The aim of this paper is to identify and describe cognitive processes occurring during the translation of a complex wordplay. Humorous texts have received little attention in empirical translation process research. This exploratory study seeks to develop a research methodology that is suited for the empirical study of wordplay. The experiment comprised the translation of a short adversarial wordplay by three groups of translation students. Their discussions were recorded and analyzed. The results indicate that students possess a high level of awareness of the importance of linguistic structures in creating the desired humorous effect, even though they cannot always translate this theoretical knowledge into practice. Another finding was that collaboration helps students overcome mental blockages and supports creative solutions.

Keywords: empirical translation process research, wordplay translation, cognitive processes, adversarial humor, structural parallelism

1. Introduction

There exists a considerable body of empirical research into the cognitive processes involved in translation, with different studies attempting to shed light on specific phenomena. Some have been concerned with modeling the translation process, providing descriptions of the various process phases (e.g. Jääskeläinen 1999), and recently, even of pauses (e.g. Immonen 2006), others have focused on translators' problem-solving strategies including the use of translation aids (e.g. Krings 1986), or addressed the issue of expertise in translation (e.g. Englund Dimitrova 2005). Several studies have been dedicated to evaluating the usefulness of research methods, such as think-aloud protocols (TAPs), dialogue protocols, keyboard logging, eye-tracking and others (for a comprehensive overview, see Krings 2005; Göpferich 2008). One aspect that has

however received limited attention to date is the impact of the source text genre on the translation process. One genre that demands a higher than average level of creativity is humorous texts. Translations of humorous texts produced by different translators are likely to diverge significantly, which makes them a particularly interesting object for the study of underlying translation processes.[1] An analysis of the processes involved in humor translation will show not only how genre characteristics influence translation, but also provide insights into what the elusive notion of "creativity" may mean for translation.

This paper summarizes the results of an experimental study that involved the translation of a short humorous text, an adversarial wordplay (also labeled "language game," see Section 3) known as "trumping" or bantering dialogue, in which the two interlocutors try to outdo each other with witty repartees. The translation was carried out by three groups of students, all translating from English into Swedish, their mother tongue. The assumption is that a close analysis of the students' reasoning as they went about the task and their problem-solving strategies will reveal characteristics of the cognitive processes at work during translation. In order to identify and describe the cognitive characteristics of the translation process, the following two questions are addressed: (1) What seems to trigger translation solutions? (2) How do the students arrive at decisions? To study the processes, and the products, i.e. the final translations, an approach that integrated concepts from cognitive linguistics and creativity research was employed.

Section 2 discusses the theoretical background to the empirical study. Section 3 provides a description of the methodological approach that was used, and Section 4 presents a summary of the results. In the concluding discussion in Section 5, the results are critically evaluated and suggestions for further study are formulated.

2. Theoretical background

Taking models from cognitive linguistics and creativity research and adapting and integrating them into methods for analyzing creativity-demanding translations is not a new approach. In one of the most influential studies, Kussmaul (2000)[2] claims that creativity is required by the translator in all cases in which direct translation is not possible. In an attempt to demystify creativity, Kussmaul argues that creativity is a natural phenomenon, which leads him to conclude that the thought patterns behind creativity can be learned by everyone, at least to a certain degree.

Of the many concepts and models from cognitive linguistics that can be put to use in Translation Studies, Kussmaul considers the concept of *focusing* (2000: 192)

1. Aspects that have been studied include the translatability of humorous texts (e.g. Vandaele 2001; Zabalbeascoa 2005).

2. The first to discuss creativity in translation was probably Wilss (1988).

to be particularly relevant, since it is crucial to a range of creativity techniques. Another pertinent concept is Fillmore's *scenes* and *frames* theory (Fillmore 1977), which Kussmaul (2000: 150ff.) uses as the basis for a comprehensive typology of creativity-techniques; these include reframing, the exchange of scene elements within a frame, exchange of scene elements within a scene, scene changes, scene enlargement, etc. In order to explain creativity in wordplay translations, Kussmaul draws on Schank's notion of Thematic Organizational Points (TOPs). TOPs link events and experiences that belong to totally unrelated scenarios or scenes. They constitute a useful notion in the study of translation processes, Kussmaul argues, because the relationship between the scenes (or between the larger scenarios) that emerge in source text (ST) wordplays and corresponding target text (TT) wordplays also tends to be tenuous and often of a formal rather than substantive nature (2000: 173).

From creativity research, Kussmaul adopts the concept of *lateral thinking* (de Bono 1970), which denotes a kind of associative thinking that goes beyond causal-logical thinking and is accessed when direct translation is not possible and the translator has to identify other frames or scene elements. Kussmaul also adopts the classic four-phase model, which has been used to describe the creativity process. In translation, Kussmaul claims, the four phases, i.e. preparation, incubation, illumination and evaluation, often seem to overlap (Kussmaul 2000: 78ff.). The incubation phase is characterized by mental blockades, which in TAPs are normally represented as pauses. The translator can try to overcome these blockades by resorting to activities that can lessen the cognitive pressure, such as getting something from the refrigerator or taking a walk. The evaluation phase plays a particularly crucial role in the translation process. In the illumination phase, translation trainees often succeed in formulating good solutions, which, however, then do not appear in the final product. The reason, according to Kussmaul, is insufficient evaluation (2000: 76).

Another aspect we will briefly look at is the impact of commitment to the task on the quality of the translation. TAP research has found that the level of engagement in the translation process has a considerable impact on the quality of the end-product (see e.g. Jääskeläinen 1999).

3. Method

In selecting a suitable text for this study, the following three criteria were considered crucial: (1) the wordplay that the students were asked to translate (from English into Swedish) should present no major comprehension problems, so as not to detract the test persons' attention from the translation task; (2) no direct translation should be possible; and (3) the translation of the content should require reframing or a scene change in order to allow the study of creativity-demanding processes.

Eventually, the following example was selected:

> Mr. and Mrs. Winston Churchill after an election defeat in 1945:
> Mrs. Churchill (consoling tone): Perhaps it is a blessing in disguise, dear.
> Mr. Churchill (angry tone): Well, it must be a bloody good disguise then.

The dialogue is from Veale, Feyaerts and Brône (2006: 313) and is an example of adversarial trumping. In the dialogue, the linguistic form of the initial utterance by the instigating speaker (S: Mrs. Churchill) is subverted by a humoristic counter-utterance from the responding hearer (H: Mr. Churchill), who trumps the speaker by demonstrating a so-called "hyper-understanding" (Veale, Feyaerts and Brône 2006: 315; see also Brône 2008) of the lexical-conceptual meaning of the initial utterance by reusing the word "disguise." For trumping to constitute a successful humor strategy, the form or content of the utterance of the instigating speaker must be reused. The hearer's affirmative "Well" at the start is significant, Veale, Feyaerts and Brône argue, since its use

> presupposes acceptance on the part of the hearer (H) of the premises on which S bases the instigating attack. Tacit in this acceptance is the idea that the hearer can defeat the speaker even using the linguistic weapons (logical premises) the speaker himself has chosen. (2006: 313)

The conclusion is that there is both a cognitive and a social component to this style of humor: "The humor […] arises out of a combination of this rather pointed cleverness and the social dynamic of one agent defeating another" (ibid.).

A signature characteristic of trumping as a humor strategy is the use of *parallelism* (Veale, Feyaerts and Brône ibid.). Veale, Feyaerts and Brône distinguish between five main classes of parallelism in adversarial "language games" (wordplay), namely homophonous, structural, metonymic, metaphoric, and referential parallelism, which are then further categorized into sub-classes (2006: 318). The relation between the utterances in the dialogue described above may be considered a *lexical parallelism* (a sub-class of metaphoric parallelism), since the word "disguise" is used literally in the trumping answer ("a bloody good disguise"), instead of metaphorically, as in the starting idiom "a blessing in disguise."[3] It expresses a specific instance of "disguise" (2006: 322).

Concerning the humorous effect of this particular dialogue, Veale, Feyaerts and Brône (2006: 313) assert, "[i]t is this ability of H to subvert the particular metaphor employed by S that makes H's response so witty." The humorous effect is further strengthened by the incongruity between, on the one hand, the semantic discordance between the words "blessing" and "bloody," and on the other hand, the formal similarity of these two words (alliteration).

It was stated above that an important criteria for the choice of text was that it should not present comprehension problems. It can be assumed that the test persons

3. The first instance of the idiom "blessing in disguise" is from 1746, according to OED, which as a definition gives "an apparent misfortune that works to the eventual good of the recipient."

had encountered the two words "blessing" and "disguise" before and that they had no major difficulties deducing the meaning of the idiom. However, as it is not a very commonly used idiom, the students probably could not immediately access a standard translation equivalent in Swedish. One of the most comprehensive contemporary English-Swedish dictionaries, *Norstedts stora engelsk-svenska ordbok*, gives only one equivalent, i.e. "tur i oturen," which literally means, "luck in bad luck." The visual image that is created by the idiom "blessing in disguise" – something positive that is hidden in something negative – is not present in the suggested Swedish translation. Further, the words "bless" and "disguise" both have richer metaphorical meanings than the Swedish equivalents "välsigna" and "förklädnad." The fact that there is no word in Swedish with the same semantic range as "disguise," which could be used in both the initial utterance and the repartee, can be expected to present a considerable challenge to the translators. If the translation is to retain its humorous effect, a significant level of translational creativity will need to be employed.[4]

The observational data were obtained from three groups of students studying translation at the universities of Uppsala and Stockholm. All the students had Swedish as their mother tongue and translated as usual from English into Swedish. They had been studying together for at least one term and could therefore be considered to know each other reasonably well. Self-brought dictionaries were the only available aid. The data comprise taped recordings of the students' discussions during the translation process that were later transcribed, all the versions they noted down as they went about the task, and the final products, all of which were analyzed. The test persons were asked to work in groups of three. The subjects were given a warm-up task, which involved the translation of another humorous dialogue, which was fairly easy to translate. They were then asked to complete the test task within the specified time (around 30 minutes). No specific brief was formulated as it was assumed that all test subjects understood that the humorous effect of the wordplay needed to be replicated in the translation. Later reviews of the recordings confirmed that retention of the humor was accorded priority status by the test persons throughout the process.

4. Results

This section describes how the three groups approached the task and the translation alternatives suggested during the process. I will also analyze the different stages in the translation process in terms of creativity phases and briefly comment on some process characteristics and on the quality of the final products.

Group 1 consisted of three female students, who usually came up with new alternative solutions after longer pauses. Suggestions tended to be encouraged by other group members using sub-word level discourse particles, such as "hm."

4. The slightly old-fashioned tone of the language may also have posed problems even though the students did not comment on this.

The following is a summary of the solutions proposed during the discussion.[5] These solutions are presented in chronological order. The last solution quoted below (and marked with an asterisk) was submitted as the final product. I have also included a comment by one of the test persons that I consider to be crucial to the translation process.[6]

> en ulv i fårakläder (a wolf in sheep's clothes)
>
> tur i oturen (luck in bad luck)
>
> en välsignelse på omvägar – vilken omväg sen då! (a blessing on a detour – what a detour then!)
>
> kanske för ditt bästa – bäst för vem? (perhaps for your best – best for whom?)
> en välsignelse i det fördolda – då måste det vara väl dolt (a blessing that is hidden – then it has to be well hidden)
>
> det kanske ändå är för det bästa – bäst för vem? (it is perhaps for the best after all – best for whom?)

At this point, one of the test persons made a significant comment:

> "Fast andemeningen är ju ändå liksom att det var för det bästa trots allt, men det blir inget roligt" (Although the basic sense is still somehow that it was for the best anyway, but it is still not funny)
>
> det kanske finns en tanke bakom det här – det kan jag inte tänka mig (there is perhaps a thought behind this – I cannot think of it)
>
> * det är kanske en välsignelse på omvägar – den omvägen måste verkligen vara lång (it is maybe a blessing on a detour – this must be a really long detour)

The members of group 1 did not seem to be very happy with their final solution, but seem to have felt that they had exhausted the range of alternative solutions that they were able to come up with at the time.

Group 2 consisted of two male students and one female student. They showed great enthusiasm for the task and laughed a lot, but also showed strong disappointment when they could not find an acceptable solution. They discussed many possible equivalents, and some of the more realistic ones are noted here, ending with their selected end-product (*):

---

5. Variants of these were also considered, but they are not noted here due to space limitations.

6. Concerning the transcription of considered solutions and other verbalizations, standard orthography was adopted in order to facilitate readability. A dash is used to indicate the border between the two utterances in the dialogue; no dash means that only the first utterance was discussed.

förklädd välsignelse (disguised blessing)

det här kanske är det bästa som nånsin har hänt – ja, jävligt bra (this is perhaps the best that has ever happened – yes, damned good)

det kanske är lika bra, älskling – det tycker jag verkligen inte (it is perhaps just as well, darling – I really do not think so)

* det finns flera fiskar i havet, älskling – fast det är val jag vill ha (there are plenty more fish in the sea, darling – but it is a whale/the election I want to have)

The members of group 2 seemed rather tired by this stage, and although they did not think that their chosen translation was very funny, they decided to submit it.

Group 3 consisted of three female students. They critically evaluated other members' suggestions, most often using succinct comments, although they also occasionally laughed. Their considered solutions are reproduced below (an asterisk will again denote their final solution), together with two central remarks uttered during the translation process by two different group members:

kanske är det en förklädd välsignelse, älskling – i så fall måste den vara jäkligt bra förklädd (perhaps it is a disguised blessing, dear – then it must be wickedly good disguised)

inget ont som inte har något gott med sig – det ska vara jävligt gott i så fall (nothing evil that does not have anything good with it – it must be something damned good in that case)

varje moln har en silverkant – jävla liten silverkant (every cloud has a silver lining – damned small lining)

efter regn kommer solsken (after rain comes sunshine)

det finns en mening med allt som sker – den måste vara jävligt ogrammatisk i så fall (there is a meaning with everything that happens – then it has to be damned ungrammatical [in Swedish the word "mening" can signify "meaning," "sense" or "sentence"])

At this point, one of the test persons reacted with this statement:

"Gillar det första där" (I like the first part)

kanske är det något som är rätt med detta, älskling – ja, rätt åt skogen (perhaps there is something right with this, darling – yes, right to the forest [i.e. "straight to hell"])

One of the test persons then stated:

"Grundmeningen här är väl att hon ser att det kan finnas nåt som är bra med detta, men han ser inte det" (The basic sense here is I suppose that she sees that there can be something good with this, but he does not see it)

kanske döljer sig något om hörnet (perhaps there is something hidden around the corner)

* det kanske finns en djupare mening med det här, älskling – då är den jävligt bra nedgrävd (there may be a deeper meaning with this, darling – then they have buried it bloody deep)

– då är det för djupt för mig, inte så att jag bottnar, då har nån grävt ned den djävligt djupt, då måste den vara jävligt djupt nedgrävd i så fall (– then it is too deep for me, not that I can reach the bottom, then somebody has dug it damned deep, then it has to be buried damned deep in that case)

The members of this group eventually rejected the alternative formulations of the punch line and seemed reasonably satisfied with the chosen final solution.

As regards the attitudes displayed by the groups during the translation process, there was a significant change from high confidence at the start, which was perhaps inspired by the unproblematic warm-up task and easy comprehensibility of the idiom "a blessing in disguise," to apprehension and concern. This is reflected in moments of silence in groups 1 and 3, and negative exclamations over their own performance in group 2.

Analyzing the students' interactions and discussions as a creativity process, the following picture emerges. As was mentioned above, the incubation phase in a creativity process is normally a time for associative thinking in which translators try to overcome mental blockages. With three persons collaborating on the translation, thought blockages were quickly overcome, so that no long silences occurred; instead, more or less continuous brainstorming prevailed. The paraphrasing that groups 1 and 3 resorted to half-way through the translation (see the transcriptions above) supported the lateral thinking process, and ultimately allowed them to visualize an appropriate new scene that, as has been pointed out in Section 2, is required for the successful translation of wordplay. In both cases the chosen strategy was successful and the two groups came up with good alternatives nearly directly after the paraphrasing, which also meant that they were able to advance to the illumination and finally the evaluation phases.

The analysis also showed that the translation process is not always a linear progression from preparation phase to incubation, illumination and final evaluation (cf. Section 3). They were thinking-aloud together, which meant that incubation, illumination and evaluation phases frequently overlapped due to reoccurring negative evaluations that required the formulation of new proposals. However, the minimal feedback, which was often expressed as a slightly surly "hm," may have only slightly dampened the test persons' enthusiasm for the task.

It was previously mentioned that TAP research found that the level of engagement in the translation process has a considerable impact on the quality of the end-product. The same cannot be said here. Group 2 showed a high level of commitment, yet this had little positive effect on the quality of their proposed solutions. The best solutions were proposed by group 3. In this group, all three participants were active

and their contributions were more equally distributed than in the two other groups. The students also displayed a high level of ambition, and frequently refused to accept appropriate proposals, and instead continued to search for (even) better alternatives.

As regards the acceptability of the students' final translations, I would argue that the solution submitted by group 1 was only barely acceptable, while the second group's proposal was definitely not acceptable. The first group managed to incorporate a lexical parallelism, which was mildly humorous, but the idiom used in the first line does not actually exist in Swedish, which makes the translation sound rather artificial. The end-product of group 2 constitutes a wordplay; however, since the proposed solution ignores the original context, it cannot be claimed that the humorous effect was kept. Concerning group 3, I would assert that the humorous effect is very well retained in their end product – "det kanske finns en djupare mening med det här, älskling – då måste den vara jävligt djupt nedgrävd i så fall" ("there may be a deeper meaning with this, darling – then it has to be buried damned deep in that case"). The punch line represents the typical trumping of an abstract noun by a concrete one (like in the ST), which here is complemented with a spatial dimension ("djupare mening – jävligt djupt nedgrävd" ["deeper meaning – buried damned deep"]). The central metaphorization in the ST of something positive hidden in something negative, which is trumped by the answer implying (with a vigorous formulation) that "any such positive perspective is so well hidden as to be virtually non-existent" (Veale, Feyaerts and Brône 2006: 313), is also replicated here, with the paralleling of "djupare/djupt" ("deeper/deep").

All three members of group 3 contributed to this solution. Test person 3-A uttered the word "mening" ("meaning") for the first time in the suggestion "det finns en mening med allt som sker" ("there is a meaning with everything that happens"), which was positively evaluated by test person 3-B (see above). Test person 3-C later returned to this word in a proposal that was accepted by the other two ("det kanske finns en djupare mening med det här, älskling" ["there may be a deeper meaning with this, darling"]). It can be argued that group 3's more dynamic collaboration helped them find a solution they all were reasonably satisfied with.

Especially in the early stages of the translation process, the groups tended to formulate solutions that closely resembled the original in their form, but that were ultimately rejected. For instance, the first solution proposed by group 1 – "en ulv i fårakläder" ("a wolf in sheep's clothes") – constitutes, like its ST counterpart, an interesting example of a comparative idiom with the preposition "in"/"i" ("blessing in disguise" – "ulv i fårakläder" ["wolf in sheep's clothes"]). The solution was potentially inspired by the formal similarity in Swedish between "förklädnad" ("disguise") and "fårakläder" ("sheep's clothes"). However, while the ST idiom implies something positive hidden in something negative, the TT proposal reverses this and stipulates something negative hidden in something positive. To what extent this fact contributed to the rejection of the suggested solution is difficult to ascertain. A more probable explanation for the rejection of this particular solution is that no ready translation for the punch line presented itself in which the words "ulv" and "fårakläder" ("wolf," "sheep's clothes") would reappear.

An interesting observation is that none of the groups felt that the first word of the punch line, "Well," needed to be translated. While it may be argued that "well" adds little content, its non-translation weakens the social impact of the trumping strategy by ignoring the initial affirmation of the speaker's utterance by the hearer (see Section 3).

5. Conclusions

The aim of this paper was to describe how three groups of students tackled a complex translation task and analyze the cognitive processes involved in translating an adversarial wordplay. The text that they were asked to translate presented major challenges that required a good measure of creativity. Although the wording of the dialogue itself was easy to comprehend, as was the idiom that was used, it was difficult to find an equivalent in Swedish that represented the same visual image as the original. "A blessing in disguise," which in the repartee is then converted into "a bloody good disguise," suggests that the instigating speaker's claim, i.e. that something positive is hidden in something negative, is interpreted by the hearer as being so well hidden that it is practically nonexistent. The equivalent suggested in Swedish dictionaries for "a blessing in disguise," i.e. "tur i oturen" (literally: "luck in bad luck"), proved impossible to use because it could not be sufficiently adapted to make sense in the punch line.

Several of the solutions discussed by the students employed structural parallelism (see Section 3). However, often only one word was repeated and not the entire structure. Group 1, for example, formulated suggestions in which the words "omväg" ("detour") and "tanke/tänka mig" ("thought/think of") were reused: "en välsignelse på omvägar – vilken omväg sen då!" ["a blessing on a detour – what a detour then!"], and "det kanske finns en tanke bakom det här – det kan jag inte tänka mig" ["there is perhaps a thought behind this – I cannot think of it"]. The resulting translations produced only a minor humorous effect, however, which was also recognized by the group. Group 3 applied a type of metaphoric parallelism (sub-class lexical parallelism), "det kanske finns en djupare mening med det här, älskling – då måste den vara jävligt djupt nedgrävd i så fall" ("there may be a deeper meaning with this, darling – then it has to be buried damned deep in that case"), which produced a significantly greater humorous effect. The image of the ST dialogue of something positive hidden in something negative was also very well replicated, which further strengthened the humorous effect.

Paraphrasing, which groups 1 and 3 employed as a strategy halfway through the translation, seemed to engender new and fresh perspectives and, consequently, led to better solutions. Group 3 had in fact cited the word "mening" ("meaning") even before they began to paraphrase, but it was only when they started to explore the usefulness of the word "grundmeningen" ("the basic sense") that they were able to construct an appropriate image or scene that could be used in translating both the instigating

utterance and the repartee. The suggestion to render the aspect of "hiding" through "nedgrävd" ("buried") was made soon afterwards.

The test subjects of this study were university students enrolled in a translation program. It would be of interest for further studies to compare their approach to the task with the strategies employed by other groups, such as professional translators and bilingual individuals with no professional experience in translation. In order to ensure a valid comparison of the cognitive processes involved in a complex translation task, it will however be important to carefully assess the level of expertise of every test person (see Englund Dimitrova 2005: 10–19).

References

de Bono, E. 1970. *Lateral Thinking. A Textbook of Creativity*. London: Ward Lock Educational.

Brône, G. 2008. "Hyper- and misunderstanding in interactional humor." *Journal of Pragmatics* 40: 2027–2061.

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process* [Benjamins Translation Library 64]. Amsterdam/Philadelphia: John Benjamins.

Fillmore, C. J. 1977. "Scenes-and-frames semantics." In *Linguistic Structures Processing*, A. Zampolli (ed.), 55–88, Amsterdam: North Holland.

Göpferich, S. 2008. *Translationsprozessforschung: Stand – Methoden – Perspektiven.* Tübingen: Gunter Narr.

Immonen, S. 2006. "Translation as a writing process. Pauses in translation versus monolingual text production." *Target* 18 (2): 313–335.

Jääskeläinen, R. 1999. *Tapping the Process. An Explorative Study of the Cognitive and Affective Factors Involved in Translating* [University of Joensuu Publications in the Humanities 22]. Joensuu: Joensuun Yliopisto.

Krings, H.-P. 1986. *Was in den Köpfen von Übersetzern vorgeht* [Tübinger Beiträge zur Linguistik 291]. Tübingen: Gunter Narr.

———. 2005. "Wege ins Labyrinth – Fragestellungen und Methoden der Übersetzungsprozessforschung im Überblick." *Meta* 50 (2): 342–358.

Kussmaul, P. 2000. *Kreatives Übersetzen* [Studien zur Translation 10]. Tübingen: Stauffenburg.

Vandaele, J. 2001. "'Si sérieux s'abstenir': Le discours sur l'humour traduit." *Target* 13 (1): 29–44.

Veale, T., Feyaerts, K. and Brône, G. 2006. "The cognitive mechanisms of adversarial humor." *Humor* 19 (3): 305–338.

Wilss, W. 1988. *Kognition und Übersetzen. Zu Theorie und Praxis der menschlichen und der maschinellen Übersetzung.* Tübingen: Niemeyer.

Zabalbeascoa, P. 2005. "Humor and translation – an interdiscipline." *Humor* 18 (2): 185–207.

"Can you ask her about chronic illnesses, diabetes and all that?"

Claudia Angelelli
San Diego State University, USA

Conversations about chronic illnesses, diseases running in the patient's family, as well as specific medical record questions are frequent in medical encounters. In the United States, as in most countries, these conversations frequently occur between culturally and linguistically diverse patients and providers, and many times they are brokered by an interpreter.

In this study I present a transcript that captures a typical healthcare-provider-patient conversation about chronic illnesses and medical history and run it through the lens of critical discourse analysis. I examine the power differentials between interlocutors and the role the interpreter plays in mitigating or reinforcing such power. I look at the co-construction of the meaning of "chronic" among the interlocutors, and the way in which the three orchestrate team efforts to de-construct and re-construct the notion of chronic illnesses exploring a life in-between countries and medical traditions. The data is part of a larger ethnographic study (Angelelli 2001 and 2004a) conducted in a public hospital. Interpreters work for Spanish-speaking patients and English-speaking healthcare providers in face-to-face and over-the-speakerphone interpreted communicative events (ICEs).

Keywords: patient/doctor communication, interpreted communicative event, role, power, solidarity

1. Introduction

Significant and thorough communication about chronic illnesses is crucial to a medical interview. The information collected during this stage of the interview is usually the basis on which most diagnoses are made and therapies are based (Gaston-Johansson, Albert, Fagan and Zimmerman 1990). Given the importance of medical history in healthcare provider/patient interaction, one would assume that providers take it upon themselves to find out the information they need. In order to accomplish

this, providers have to retain control of the interview and exercise the power granted to them by their position in the interaction. Occasionally this is done by observing the principle of relevancy (Heath 1992) and not allowing "digressive" narrative to be included during the interview.

Communication about patients' medical history can occur in medical encounters where the healthcare provider and the patient speak the same language, and still be a challenge for both patients and healthcare providers (Shuy 1976). Often, patients cannot discriminate between relevant and less relevant information, or simply do not understand the task at hand. Other times, patients may embark upon a lengthy and serious narrative about certain medical conditions, but their facial expressions and body language may be inconsistent with their claimed condition, leaving it unclear to the healthcare provider whether the narrative is a true representation of the patient's condition (Sternbach 1981; Lipton and Marbach 1984).

These struggles become magnified when provider and patient do not belong to the same culture and do not speak the same language. Language and culture are bound together in the communication of pain and illnesses, and cultural background is one of the major factors determining how people communicate and express issues about their health (Sternbach 1981; Lipton and Marbach 1984). Culture has been found to influence the reporting of pain and illnesses (DeTrill and Kovalcik 1997). Culture is a significant force in shaping beliefs and behaviors, giving meaning to the pain experience and influencing the way it is expressed. Illnesses are also bound to cultural ways of conceptualizing, treating and discussing them. Non-verbal communication also differs between cultures (Davidhizar, Shearer and Giger 1997). Additionally, family values and the cultural life of the community may influence reports by children on pain and illnesses. In their communication about pain and illnesses, children may reproduce the cultural views and societal blueprints of the society to which they belong. When in addition to cultural differences linguistic barriers exist that prevent health providers from understanding what the patient says, a medical interpreter is called upon to bridge these barriers.

In this paper I examine an interpreted communicative event through the lens of critical discourse analysis (CDA). I present segments of a medical interview in which chronic illnesses are explored. Using CDA I look at the power differentials between interlocutors and how the interpreter mitigates or reinforces power differentials causing (or not causing) power shifts, perpetuating status quo and/or exercising other strategies. I look at the co-construction of the meaning of "chronic" among the interlocutors, and the way in which the three participants orchestrate team efforts to de-construct and re-construct the notion of chronic illnesses throughout the patient's life in-between countries and medical traditions. First I proceed to review relevant literature and contextualize the study. I then examine the data though the lens of critical discourse analysis, before drawing preliminary conclusions.

2. Review of relevant literature

Like any monolingual medical consultation, an interpreted communicative event (ICE) could also be divided into the six phases that Byrne and Long (1976, in Heath 1992: 237) identified in their study of verbal behavior of 2000 medical consultations. Those phases are: (1) relating to the patient, (2) discovering the reason for attendance, (3) conducting a verbal or physical examination or both, (4) consideration of the patient's condition, (5) detailing treatment or further investigation, and (6) terminating. Contrary to a monolingual medical consultation, in which power imbalances are generally stable and salient (Heath 1992), in an ICE there are three participants, and the presence of the interpreter does impact the power differential as well as all six phases of the medical interview. The impact is not only on the type and amount of information that gets communicated, but also on how regular power/solidarity patterns get altered or reinforced. From previous research in communication we know that shifts occur in interactions when, for example, interlocutors engage in expressing power and/or solidarity (Tannen 1984). Specifically when we look at interactions mediated by interpreters we see that interpreters' renditions vary according to intention. Wadensjö's taxonomy (1998) of interpreters' linguistic behavior, later expanded by Angelelli's discussion on interpreter's agency (2004a), offers a solid starting point to discuss power, solidarity and power shifts during the medical interview.

Previous studies in interpreted-mediated consultation show examples of interpreters' agency (Angelelli 2001, 2004a and 2004b) materialized through the following behaviors: editing by omitting or adding (Davidson 1998; Wadensjö 1998; Angelelli 2004a), interpreter as principal interlocutor (Roy 2000; Angelelli 2004a), interpreter as facilitating understanding (Davidson 1998; Angelelli 2004a), interpreter modifying register (Angelelli 2004a), and interpreter exercising agency by being visible in the talk (Angelelli 2004a). Interpreters play key roles in perpetuating power differentials or in causing shifts in it. Interpreters may cause power shifts by exercising their agency. Interpreter's agency manifests itself, for example, when interpreters do one or more of the following:

1. They introduce or position the self as a party to the ICE. In this way they become co-participants (Wadensjö 1998; Roy 2000; Angelelli 2001, 2003 and 2004a) and co-constructionist (Davidson 1998);
2. They set communication rules (e.g. turn-taking) and control the traffic of information (Roy 2000; Angelelli 2001, 2003 and 2004a);
3. They paraphrase or explain terms or concepts (Davidson 1998; Angelelli 2004a);
4. They slide the message up and down the register scale (Angelelli 2004a);
5. They filter information (Davidson 1998; Angelelli 2004a);
6. They align with one of the parties (Wadensjö 1998; Angelelli 2004a); and
7. They replace one of the parties to the ICE (Roy 2000; Angelelli 2004a).

These previous studies on the role of the interpreter and the analysis of the information transferred during a medical interview have used lenses other than critical discourse. In this paper the transcript is read through the lens of critical discourse.

3. The study[1]

3.1 Procedures

The data analyzed in this paper is part of a larger ethnographic study (Angelelli 2004a) conducted in the United States. During 22 months I observed ten Spanish/English medical interpreters at California Hope.[2] I recorded, transcribed and analyzed 392 interpreted-communicative events (ICEs). During the ethnographic study, in addition to collecting the 392 ICEs, I conducted participant and non-participant observations, took notes, interviewed the interpreters and the manager of Interpreting Services using semi-structured protocols, administered the Interpreter Interpersonal Role Inventory (Angelelli 2004b) and collected various artifacts.

At California Hope, ICEs occurred both face-to-face and over the speakerphone. For face-to-face encounters, interpreters walked to clinics and wards. For ICEs conducted over the phone, they stayed in their workstation and took the calls. In the case of face-to-face ICEs, I was present during the interaction, observing, taking notes and audio-recording. When ICEs occurred over the phone, I sat with the interpreter in the workstation and plugged my tape recorder into the interpreter's telephone. Participation in the study was voluntary, all clearances for protection of human subjects in research studies were pursued, and informed consent was obtained for each ICE. Consent for recording and observing was requested immediately before the interpreting started. During speakerphone interactions, the parties (healthcare provider and patient) were requested to consent orally, and that was the time when they were alerted to the presence of a researcher in the interaction. Sometimes the request entailed just asking for consent. Other times the participants engaged in a conversation about the nature of the study.

3.2 The site

California Hope is a public hospital on the West Coast of the United States. At the time the larger study was conducted, California Hope was the first hospital in the

1. For further details about the study in general, and specifically about the site, the participants and interpreting via speakerphone, the reader is directed to Angelelli (2004a).

2. All names of individuals, organizations and geographical areas are fictitious to protect identities of participants and sites.

country to offer on-site telephone interpreting as well as face-to-face. To this date it is the only hospital in the county with an open door policy that guarantees access to needed medical care, regardless of patients' ability to pay. It enjoys a strong reputation as the community's healthcare provider. The population served by California Hope ranges from middle class to working class, and the average patient lives below the poverty line. The community serviced by California Hope is made up of a diverse ethnic pool. Most of the patients are African-American, Asian or Hispanic, though in terms of interpreting, Hispanics constitute the biggest group (26.6%) served by Interpreting Services at California Hope.

3.3 The participants

Ten full time Spanish-speaking interpreters plus two who work as extra help or part time, and one who works during night shifts at the Emergency Room, participated in the larger study. Vicente,[3] the interpreter who we will encounter in this paper, is one of the ten full-time Spanish-speaking interpreters who work between 9:00 a.m. and 5:00 p.m. Monday through Friday. Vicente was born in Peru and brought up in Romania, where he pursued higher education in medicine. He graduated as a general physician and later came to live in the United States. At the time the study was conducted he was 60. Vicente has worked as a full-time staff interpreter at California Hope for 15 years. To join the team of interpreters at California Hope Vicente met the following requirements: two years experience in the field (as medical interpreter, translator, bilingual medical assistant, etc.), bilingual ability and passed the Interpreting Services test. The test measures medical vocabulary, interpreting skills and memory retention in both directions (English into Spanish and Spanish into English). The test progresses from short phrases to complex paragraphs. The test was designed as an in-house instrument and at the time the study was conducted no information on validity or reliability studies on the test was available. Like all Spanish-English interpreters working at California Hope, Vicente brokers communication between Spanish-speaking patients (mostly working-class immigrants) and English-speaking healthcare providers.

3.4 Data collection and coding

Observing the principles of ethnographic study, I had spent a significant amount of time before I started recording getting acquainted with the site, shadowing interpreters and building trust. By the time the data collection started I had already shadowed and monitored Vicente (as well as the other Spanish interpreters) for over 9 months. I arrived at Interpreting Services at 8:30 and spend the day with Vicente two to three

3. All names are pseudonyms in order to protect the identity of the participants in the study.

Table 1. Inventory of total ICEs according to modality and Vicente's share

Total of ICEs	**392**	**%**	**% of total for Vicente (6%)**
Face-to-face	11	2.80	9
Speakerphone	381	97.19	6.3

Table 2. Inventory of total ICEs according to purpose and Vicente's share

Total of ICEs/purpose	**392**	**%**	**% of total for Vicente (6%)**
Make an appointment	120	30.61	5.8
Place a phone call	10	2.55	0
Give results/news	4	1.02	0
Educational session	9	2.29	22
Visit/check	230	5.86	5.6
Other (complain about bills, etc.)	19	4.84	15.8

times per week. I hooked a headset to his telephone and walked with him to face-to-face interviews, and he interacted with me as he did with other new interpreters in training.

ICEs vary widely at California Hope in terms of length, time of the day, and so forth. Tables 1 and 2 illustrate the variety of the 392 ICEs collected for the larger study (Angelelli 2001 and 2004a) for ten interpreters and the share produced by Vicente.

At California Hope 392 ICEs occurred in two different modalities: face-to-face or remotely (i.e. over the speakerphone). In face-to-face, the three parties (i.e. patient, healthcare provider and interpreter) interact face-to-face. In the modality called "speakerphone," patient and healthcare provider are face-to-face and share a space in which there is a speakerphone. The interpreter interacts remotely. Using a telephone connected to the speakerphone, the interpreter works from a remote location.

The ICEs produced by Vicente varied from 2 to 90 minutes in length. He interpreted interactions face-to-face and over the speakerphone. The nature of the medical visits varied from admissions to discharging, pre-natal sessions to terminal diagnosis and life-end decisions, from general medicine to specialized treatment and from speech therapy to mental ward hearings. The ICEs produced by Vicente (as well as for the other nine interpreters) were first indexed, before the index entries were used for unit analysis. The total number of units found was 392 ICEs, of which Vicente produced twenty-five (6.4%).

The analysis of the data for Vicente (as well as for the other nine interpreters who took part in the larger study) was carried out on two separate occasions using unmarked index copies (transcripts), in order to insure consistency in the identification procedure. Annotations (based on type of interaction according to modality, purpose, role of interpreter, etc.) were made in pencil for each unit (transcript of an interpreted communicative event) to allow for correction, based on category fluctuation or solidification. These annotations (codes) constituted the basis of subsequent

emerging categories and patterns (LeCompte and Schensul 1999). The pattern of the nature of the ICE generated two categories: face-to-face and over-the-speakerphone. This category was objective and there was no need for corroboration by the informant (in this case, Vicente). Another copy was used for the intention/purpose of the interaction. Six categories emerged from patterns of the intention of the ICEs, and those are shown in Table 2: make an appointment, place a phone call, deliver news/result, visit/check, educational session, and other/miscellaneous. After analyzing the ICEs according to their nature and attempting a first classification according to the intention of the interaction, I interviewed Vicente. I asked him about the purposes of the interpretations for which he is called (the types of interactions he brokers). I took notes of his answers. Then I offered him the categorization of interactions based on purposes that I obtained from the data analyzed and got his feedback. In this way I made sure that the etic perspective that I had initially gained as researcher coincided with the informants' perspective, thus becoming an emic one (LeCompte and Schensul 1999: 221). It also served as corroboration for the category of type of interactions based on purpose.

Another clean copy of the transcripts was used for the last pattern that emerged in terms of the interpreter's intervention or lack of intervention during the ICEs and its impact on maintaining or altering power differentials between doctor and patient. Those interventions or behaviors were first coded on the margin. Those codes stated characteristics of the interpersonal role of the interpreter (Angelelli 2004b) that transpired from the interaction across interpreters and across the nature and intention of the ICEs I analyzed.

In this chapter I present and analyze segments of an ICE from Vicente in order to study if/how the interpreter affects the power differentials by mitigating or reinforcing expected behaviors. From the whole corpus of interactions dealing with chronic illnesses that I collected at California Hope (including those produced by Vicente), I chose one example to discuss power management in interpreter-mediated doctor-patient consultation.

4. A look at the data through the lens of critical discourse analysis

4.1 What is critical discourse analysis?

In Van Leuven's terms critical discourse analysis (CDA) is or should be concerned with discourse as the instrument of power and control as well as with discourse as the instrument of social construction of reality (2002: 166). CDA is not a method of analysis but rather an approach (Toolan 2002). It implies a bi-directional causal relation between paired terms, such as language production of speakers during an interaction and status of speakers in that same interaction. For example, when discussing communication in a medical setting, we can say that the language produced during a

doctor-patient interview can create the resulting power differentials between doctor and patient.

At the same time, the power differentials between doctors and patients and their pre-existing assumptions about their respective roles and power or standing may be shaped by the kind of language produced by each of the parties (Prince 1986). This means for example that in order to exercise power, the doctor may use medicalese, take longer turns, and even corner the patients with the use (and sometimes abuse) of yes-no questions (Angelelli 2004a). In turn the patient may defer to the doctor, speak in short sentences, wait to be given a turn and struggle to produce the yes-no answer required by the doctor. In addition the use of language of each of these two interlocutors (patient and doctor) may be shaped by their positions of higher and lower status within the interaction and, at the same time, their positions and roles may color and shape the language they produce (Fairclough 2002).

According to Fairclough and van Dijk, an additional characteristic of CDA is that the object of study and "the phenomena in need of scrutiny are in some respects masked or covert, or embedded, or made so ordinary as to escape attention or critique" (in Toolan 2002: xxii). In this sense, CDA could be conceived as an awareness-raising device or mechanism that alerts us of the anti-democratizing effects of the discourses we use and encounter every day. CDA also allows us to see the interplay between a linguistically-oriented text analysis (whether the text is written or oral) and an analysis of the social aspects of discourse that draws upon social theory and uses discourse analysis as a form of social analysis (Fairclough 1992). As Toolan (2002) established, the emergence and development of critical discourse analysis as an analytical approach is a direct result of the tendencies of our times. One of the characteristics of our current "global" village is that it is to some degree a "textual" village. Text gets produced and reproduced through communication. As a result we swim in a sea of texts and discourses. At the same time, democracies empower us to challenge those texts and discourses, to be less trusting, to question them instead of merely accepting them. In Toolan's view, it is the intersection of the "textualist" turn and the "democratizing" turn that underpin all studies that fall under CDA (2002: xii).

In relation to the medical setting, some proponents of CDA (e.g. Wodak 2002) state that research on healthcare provider–patient communication has been of two kinds: (1) either focusing on the institution, and that was labeled the medical-sociological approach; or (2) focusing on the micro-structural aspects of communication, and that would be the linguistics approach (e.g. discourse analysis, conversational analysis, etc.). Wodak goes on to claim that seldom have those two approaches been combined into a single study and criticizes American research for being focused on the analysis of an individual conversational phenomena (e.g. question-answer sequences, narratives). This study aims to begin filling this gap by looking at the language used within an institution resulting from both the constraints placed by the institutions on the interlocutors as well as from the interlocutors' decisions to challenge or perpetuate the status quo (imposed by the institution and society) as expressed by language. I examine how power and solidarity are imposed by institutions and resisted or

perpetuated by individuals as evident in their discourse and how the discourse used creates, perpetuates or shifts power differentials during the medical interview. As stated in the introduction, in this study I have used CDA as the lens to look at the data.

4.2 The transcript

The following segments[4] come from an interview between Señora Mesa (P), a Mexican Spanish-speaking working-class elderly female patient, and a middle-aged female Caucasian nurse at California Hope (N). Vicente (I), a staff interpreter who is working remotely and through the speakerphone, facilitates the communication between healthcare provider and patient, who are face-to-face in an office. The interview is fifteen minutes long. In the tradition of ethnographic studies, the transcript was chosen based on typicality. Interactions over the speakerphone were more frequent than face-to-face ones. In addition, at California Hope the question about chronic illnesses emerged every time medical records were discussed during the history-taking phase. Like the other nine interpreters, Vicente had several instances of brokering the term "chronic illnesses." We join the interlocutors during the history-taking phase of the interview. The nurse wants to know whether the patient suffers from any chronic illnesses. She does not pose the full question for Vicente to interpret. Rather, she asks Vicente to produce the line of questioning (turn 1). In accepting the "invitation" and carrying out the nurse's request, Vicente adopts a series of strategies that are consequential to the power shifts that occur during the interview. Let's take a look:

Tape 17[5] Call 3: 200–224 (Angelelli 2004a: 94–96). This interaction has a total of 37 turns and it has been segmented into three sections for easier reference.

Segment 1

1 N Can you ask her about chronic illnesses, diabetes… all that?
2► I *Ahá. Señora Mesa, ¿alguna vez le dijo el doctor, aunque sea veinte años atrás, aquí o allá, que tenía usted diabetes*?
(OK. Mrs. Mesa, has a doctor ever told you, even twenty years ago, here or there that you had diabetes?)
3 P No.
4► I *¿Que tenía la presión alta?*
(That you had high blood pressure?)
5 P No.
6► I *¿Que tenía alguna enfermedad al corazón?*
(That you had heart disease?)
7 P Noooo.

4. The level of detail in the transcripts is adequate for the kind of analysis performed.

5. Name of interpreter and number of tape collected (for more details, see larger study).

8► I *¿Que era enferma del hígado?, ¿De los riñones? ¿Del estómago?*
(That you suffered from liver problems? Kidney problems? Stomach problems?
9 P No.

On turn 1 we see the shift of the control of the interview, from the nurse to the interpreter. The nurse is telling the interpreter what to do, not asking the patient herself. In spite of all the risks involved in relinquishing control in a medical interview, the nurse does not hesitate when she requests that Vicente produce the line of questions about chronic illnesses. Vicente does not stand up for what an interpreter should or should not do, according to statements on code of ethics or prescriptive norms of his profession. He chooses instead to follow the instructions of the nurse. One may interpret this as an instance of the interpreter preserving the status quo that places the healthcare provider at the top of the power hierarchy and recognizing that status. However, Vicente takes this opportunity to modify what is being asked of him (and of the patient) and exercises his agency (or power) to change it without perhaps being fully aware of the consequences of this behavior. Vicente could have simply asked the question literally (do you have any chronic illnesses), or perhaps modify it a little by changing the technical term "chronic" to another term, such as "long time" or "permanent," with which the patient could have been more familiar. Instead Vicente, by accepting the nurse's invitation, takes it upon himself to embark on a discovery expedition in which he considers not only illnesses but also the patient's geographic displacement and the passing of time (turn 2). In addition he decides to unpack the term chronic illnesses by displaying some (limited) examples starting with diabetes (turn 2), high blood pressure (turn 4), heart disease (turn 6) and exploring liver, kidney and stomach problems.

One could argue that the term chronic illnesses encompasses much more than these six issues mentioned, and/or that calling the patient's attention to these six issues to the detriment of others could be problematic. One could also argue that illustrating what the term chronic illness means by breaking it into its parts is better than leaving it as a technical term that may be obscure to the patient. One may also say that giving six examples of chronic illnesses is better than none. One thing is certain: in all the instances in the whole corpus where the question on chronic illnesses was asked directly by the healthcare personnel (approximately 32%), it was phrased as a closed question (yes/no answer) and addressed directly to the patient or to the interpreter. The question was never broken into smaller components. Examples of those questions are, "Do you have any chronic illnesses?", "Has she had any chronic illnesses?", "Ask her about chronic illnesses," or "Any chronic illnesses run in the family?".

In Segment 1, we see that Vicente took nine turns (actually nineteen, as we see below when the interaction continues) to ask about chronic illnesses. Every time that he offered an example of a chronic illness he checked for patient comprehension by waiting for the patient to respond. A general yes/no question thus becomes as many questions as the interpreter/patient is ready to construct through examples, expansions and clarifications (Wadensjö 1998). And as the following segment (Segment 2)

will continue to illustrate, Vicente's brokering of the term "chronic illnesses" not only covered illnesses but also surgery, hospitalizations and medicine intake.

Segment 2

10▸ I *¿Alguna vez la operaron, la internaron? ¿Allá o aquí?*
(Have you ever been operated on or hospitalized? Here or there?)
11 P Noooo.
12▸ I *¿Nunca ha estado usted enferma?*
(You have never been sick?)
13 P *Esteeee si estuve enferma pero … es deeee depresión nerviosa, no de otra cosa*
(Well … I was sick, but … it was de … nervous depression … I did not suffer from anything else.)
14▸ I *Okay … ¿y la internaron por la depresión nerviosa?*
(Okay, and were you hospitalized for nervous depression?)
15 P *Sí …*
(Yes.)
16▸ I *¿Allá o aquí?*
(Over there or here?)
17 P *Esteee … emmm … en Azusa …*
(Hmmm … in Azusa.)
18 I *¿En dónde?*
(Where?)
19 P *Azusa … cerca de Los Angeles.*
(In Azusa, near Los Angeles.)
20▸ I Ya … she is saying that she denies diabetes, denies cardiovascular disease, denies blood pressure, denies eh … problems with her stomach and her liver … she said that she was … denies surgery … she was admitted once eh … close to Los Angeles … ehmm … for depression.
21 N Okay … but she does not take any medicine now?
22▸ I *Señora, ¿está usted tomando alguna medicina estos días?*
(Ma'am, are you taking any medicine these days?)
23 P No.
24▸ I *¿Alguna medicina que compró sin receta?*
(Any over-the-counter medicine?)
25 P No.
26▸ I *¿Que trajo de allá, que le dio la comadre Juana?*
(That you brought from there, that Comadre[6] Juana gave to you?)
27 P *No, nada.*
(No, nothing.)
28▸ I Negative … negative

6. Term used in Spanish to refer to the relationship between mother and godmother of a child.

Vicente continues to expand the rendering of "chronic illnesses" by exploring hospitalization (turn 10) which the patient denies (turn 11). When it appears that he is almost done with the exploration of illnesses, he decides to verify one more time that Señora Mesa has "never being sick" (turn 12). This is when Señora Mesa mentions being sick for the first time (13). Vicente stays on that lead and asks if that depression caused her to be hospitalized (turn 14) and where the hospitalization occurred (turn 16), once again accounting for the patient's geographic displacement (16–19). After a total of 20 turns Vicente reports back to the nurse by summarizing his interaction with Señora Mesa. Vicente does render to the nurse the details of how he and Señora Mesa verified the information, clarified the location, etc. He does not report back on all the questions asked (e.g. question about kidneys). We know from previous research in dialogic interpreting (Wadensjö 1998; Angelelli 2004a) and in co-construction of meaning that, in interpreting talk cannot be treated as text because it is an activity. In addition, the point that concerns us here is how Vicente, the nurse and Señora Mesa deal with power. Vicente and Señora Mesa co-construct the notion of "chronic illness" so far (rather than focusing on the fact that Vicente omitted to report on the kidney question and if that omission is a mistake) and how the nurse relinquishes any control over the line of questioning. The nurse acknowledges receipt of Vicente's summary, takes it at face value and moves to inquire about current medicine intake (turn 21). This power shift from the nurse to the interpreter happens twice in this segment as Vicente expands on the question of medicine intake by accounting for over-the-counter medicine (turn 24), folk medicine and geographic displacement (turn 26), and again reports back only a summary (turn 28), which the nurse again takes at face value. On the basis of these exchanges and the physical exam that she is performing as the interaction goes on, the nurse assesses the situation in the final part of the interview as we see below. Since she needs supplies to perform a test, she walks out of the room, and that is when the last part of this interview ends.

Segment 3

29 N Okay, all right … sounds just like she has the blues … her lungs are clear, she is breathing fine and her color looks good. I'm gonna check her oxygen saturation and then probably send her home with advice and give her the number to call …

30 I *Okay … Señora dice nuestra enfer* [
[
(Okay … our nurse says

31► N [I'll be right back (*nurse walks out of the room*)

32 I [*mera que le ha escuchado los pulmones y suenan bien …*
(that she has heard your lungs and they are fine…)

33► P *Síííí.*
(Yeees.)

34 I *Dice que su color está bien … que … le va a medir el oxígeno de la sangre*
(She says that your skin color looks fine … that … and that she will measure the oxygen in your blood.)
35► P Okay.
36► I *Eso no duele … le va a poner una lucecita en uno de sus dedos solamente por un minuto y después la manda a su casa con unas recomendaciones.*
(That doesn't hurt … and she will put a little light in one of your fingers during one minute and then she will send you home with advice.)
37 P Okay.

On turn 29 the nurse takes over the control of the interview as she recaptures the situation and tells Vicente what she is going to do with Señora Mesa. It should be noted that the nurse talks about Señora Mesa with Vicente rather than talking to Señora Mesa directly with Vicente's help. As Vicente starts interpreting this information to Señora Mesa (turn 30), the nurse walks out of the room (turn 31) to get the supplies she needs to perform the oxygen saturation test. Vicente, now alone in the room with the patient, continues rendering the information to Señora Mesa. By accepting this situation, he not only continues to render the information produced by the nurse but he also takes it upon himself to explain to the patient how the oxygen saturation test works and reassures her that it will not be painful. The interview comes to an end.

5. Discussion and conclusions

In the data presented we see how Vicente perpetuates, shifts, uses (and even abuses) the power generally held by the healthcare provider in a medical interview. He uses different strategies to accomplish communicative goals. On the basis of his perception of the patient, he slides messages down the register scale, investigates answers, expands and summarizes utterances, and brokers cultural references. Vicente applies these strategies without a specific direction, at the prompt "and all that." Vicente's exercises power as he orchestrates the talk. During subsequent interviews with Vicente I asked him about the strategies identified in the transcripts, such as expansions (Wadensjö 1998) and register shifts (Angelelli 2004a), and why he had used them. He explained to me that many times patients either do not understand the healthcare provider's question or do not give the complete answer. He is convinced that his job involves brokering the questions in a way that patients will understand, and also to continue digging until he obtains the information that is necessary to answer the question. This transcript allows us to see how the interpreter may play a pivotal role in altering or perpetuating the power differentials between patients and doctors. Unlike what statements found in code of ethics generally prescribe, descriptive research (Wadensjö 1998; Metzger 1999; Bolden 2000; Roy 2000; Angelelli 2004a, 2004b; Meyer, Bührig, Kliche and Pawlack 2010) shows that interpreters impact the nature of the interaction. They are present (Wadensjö 1998; Roy 2000), and they both

acknowledge that they have agency and exercise such agency (Angelelli 2004a, 2004b). In this study we see how, by exercising his power, Vicente alters the power dynamics of the medical encounter. Sometimes he aligns his power with that of the provider and/or he becomes even more powerful than the provider (turn 36). Other times, he takes the power and uses it to alter the status quo by digging for information and thus allowing the patient's stories to be heard (turn 10–27).

6. Implications and call for further research

This study has practical and theoretical implications for the field of Interpreting Studies in general and for the education of healthcare interpreters and healthcare providers in particular. It also has several implications for a theory of patient-centeredness and for policy regulating equal access to services for limited-English-speaking patients. Healthcare administrators, managers and providers need to be aware of power shifts that occur when they relinquish control of the interview. In addition, healthcare providers need to account for differences when they communicate with patients who, unlike fellow healthcare providers, do not share medicalese and represent other cultures and traditions. This study also has implications for a more encompassing theory of interpreting currently focusing mostly on cognitive and linguistic issues. The agency of the interpreter cannot be denied, but should instead be problematized and accounted for. And last but not least, this study has practical and theoretical implications for the education of healthcare providers and interpreters. Healthcare providers need to learn to communicate with the patient with the help of an interpreter. They also have to reflect upon the risks present in relinquishing control of the medical interview and letting control be taken up by an interpreter who does not necessarily share their expertise. Interpreters also need to reflect upon the consequences of taking up certain responsibilities that go beyond their field of expertise.

In addition, this study offers evidence of behaviors of so called "dual-role interpreters" (CHIA 2002) that so far have not been studied. Vicente was a physician in his home country, and as shown in the transcript, he could not take off his physician's hat and put on his interpreter's hat: he was wearing both. Further studies are needed to gain a better understanding of how an interpreter's agency may or may not be used to shift power, and if the power shift occurs in one direction only (i.e. favoring one of the participants) during the whole interview or if it varies across moments in the conversation.

References

Angelelli, C. 2001. "Deconstructing the invisible interpreter: A critical study of the interpersonal role of the interpreter in a cross-cultural/linguistic communicative event." (Doctoral dissertation, Stanford University, Stanford, USA.)

———. 2003. "The visible collaborator: Interpreter intervention in doctor/patient encounters." In *From Topic Boundaries to Omission: New Research on Interpretation*, M. Metzger, S. Collins, V. Dively and R. Shaw (eds), 3–25. Washington, DC: Gallaudet University Press.

———. 2004a. *Medical Interpreting and Cross-cultural Communications*. Cambridge, UK: Cambridge University Press.

———. 2004b. *Revisiting the Interpreter's Role: A Study of Conference, Court and Medical Interpreters in Canada, Mexico, and the United States.* Amsterdam/Philadelphia: John Benjamins.

Bolden, G. 2000. "Toward understanding practices of medical interpreting: Interpreters' involvement in history taking." *Discourse Studies* 2 (4): 387–419.

CHIA Standards and Certification Committee. 2002. *California Standards for Healthcare Interpreters: Ethical Principles, Protocols, and Guidance on Roles & Intervention.* Los Angeles, CA: The California Endowment.

Davidhizar, R., Shearer, R. and Giger, J. N. 1997. "Pain and the culturally diverse patient." *Today's Surgical Nurse* 6: 36–43.

Davidson, B. 1998. "Interpreting medical discourse: A study of cross-linguistic communication in the hospital clinic." (Doctoral dissertation, Stanford University, Stanford, USA.)

DeTrill, M. and Kovalcik, R. 1997. "The child with cancer. Influence of culture on truth-telling and patient care." *Annals of the New York Academy of Sciences* 809: 197–210.

Fairclough, N. 1992. *Discourse and Social Change*. Cambridge: Polity Press.

———. 2002. "Register, power and socio-semantic change." In *Critical Discourse Analysis. Critical Concepts in Linguistics. Volume I,* M. Toolan (ed.), 304–321. New York: Routledge.

Gaston-Johansson, F., Albert, M., Fagan, E. and Zimmerman, L. 1990. "Similarities in pain descriptions of four different ethnic-culture groups." *Journal of Pain and Symptom Management* 5 (2): 94–100.

Heath, C. 1992. "The delivery and reception of diagnosis in the general-practice consultation." In *Talk at Work. Interaction in Institutional Settings,* P. Drew and J. Heritage (eds), 235–267. Cambridge: Cambridge University Press.

LeCompte, M. and Schensul, J. 1999. *Analyzing & Interpreting Ethnographic Data.* Vol. 5. Walnut Creek/London/New Delhi: Altamira.

Lipton, J. A. and Marbach, J. J. 1984. "Ethnicity and the pain experience." *Social Science and Medicine* 19 (12): 1279–1298.

Metzger, M. 1999. *Sign Language Interpreting: Deconstructing the Myth of Neutrality.* Washington, DC: Gallaudet University Press.

Meyer, B., Bührig, K., Kliche, O. and Pawlack, B. 2010. "Nurses as interpreters? Aspects of interpreter training for bilingual medical employees." In *Multilingualism at Work: From Policies to Practices in Public, Medical and Business Settings,* B. Meyer and B. Apfelbaum (eds), 163–184. Amsterdam/Philadelphia: John Benjamins.

Prince, C. 1986. "Hablando con el doctor: Communication problems between doctors and their Spanish-speaking patients." (Doctoral dissertation, Stanford University, Stanford, USA.)

Roy, C. 2000. *Interpreting as a Discourse Process.* New York/Oxford: Oxford University Press.

Shuy, R. 1976. "The medical interview: problems in communication." *Primary Care* 3: 365–386.

Sternbach, R. A. 1981. "Chronic pain as a disease entity." *Triangle* 20: 27–32.

Tannen, D. 1984. *Conversational Style: Analyzing Talk among Friends.* Norwood, N.J.: Ablex.

Toolan, M. (ed.). 2002. *Critical Discourse Analysis. Critical Concepts in Linguistics.* Volume I, xxi–xxvi. New York: Routledge.

Van Leuven, T. 2002. "Genre and field in critical discourse analysis. A synopsis." In *Critical Discourse Analysis. Critical Concepts in Linguistics. Volume II,* M. Toolan (ed.), 166–199. New York: Routledge.

Wadensjö, C. 1998. *Interpreting as Interaction*. London/New York: Longman.

Wodak, R. 2002. "Critical discourse analysis and the study of doctor-patient interaction." In *Critical Discourse Analysis. Critical Concepts in Linguistics. Volume II,* M. Toolan (ed.), 340–364. New York: Routledge.

PART III

Studies of interpreting and translation expertise

Effects of linguistic complexity on expert processing during simultaneous interpreting

Adelina Hild
CRIT, SUNY Binghamton, USA

Processing models of simultaneous interpreting (SI) have predicted that the linguistic complexity of the source-language text will adversely affect performance quality due to the increased storage and processing demands imposed by linguistically complicated constructions on a limited-capacity working memory (WM). The present paper reports an experimental expert-novice comparison study which examines this hypothesis by analysing the effect of four complexity metrics, validated in psycholinguistic research on interpreting performance: developmental level, amount of embedding, type of embedding and clause type. The results show a consistent effect of complexity metrics on novice performance, while for experts the effect was modulated by the overall level of textual redundancy. The findings suggest that syntactic complexity, as indexed by the measures used in this study, does not predict experts' accuracy in a significant way. In line with Ericsson and Kintch's theory of skilled WM, this indicates the somewhat diminished role of surface structure cues in skilled SI processing and suggests a strong component of strategic and macro processing in expert SI performance.

Keywords: syntactic complexity metrics, expert simultaneous interpreting, text redundancy, cognitive resource, skilled memory

1. Introduction

Effective management of cognitive resources has been traditionally viewed as the fulcrum of skilled simultaneous interpreting (SI). Early experimental research on SI cognition has focussed on detecting environmental and input factors contributing to cognitive load, such as input rate (Gerver 1971; Barik 1975) and white noise (Gerver 1974). Systemic differences between source and target language have also traditionally been viewed as a significant source of difficulty, and available empirical research has primarily focussed on contrastive analysis of specific syntactic constructions

taken from a SI corpora (cf. Uchiyama 1991 for Japanese and English; Setton 1998 for Mandarin and English; Seeber 2005 for German and English). The research reported in this chapter considers the effect of increased linguistic complexity on SI processing by examining skill-related variations in performance as a function of variations in syntactic complexity.

The chapter is organised as follows. Section 2 below problematises the notion of linguistic complexity in relation to models of WM and cognitive resource management. It also reviews different proposals for complexity metrics and research concerning their experimental validation in an attempt to select reliable metrics for evaluation in SI research. Section 3 introduces an individual differences study of syntactic complexity in SI, followed by a discussion of the findings and their implication in the final fourth section.

2. The concept of linguistic complexity, its operationalisation and experimental measurement

The study of linguistic complexity has been resumed in recent years with renewed vigour, motivated by new observational, developmental and experimental evidence collected since the 1980s. The linguistic approach to complexity encompasses a series of typological, cross-linguistic studies purporting to analyse global morphological and syntactic diversity (Sampson, Gil and Trudgill 2009); or to define the foundations of grammar in relation to the exigencies of our cognitive system (Hawkins 2004). Psycholinguistic approaches focus on the measurement of complexity and its effect on human processing as well as the assessment of the psychometric efficacy and diagnostic utility of existing metrics for healthy and clinical individuals (Lyons, Kemper, LaBarge, Ferraro, Balota and Storandt 1994; Roark, Mitchell and Hollingshead 2007). The third approach that enriches complexity research comes from the natural language processing community and deals with the possibility of automatically extracting measures of syntactic complexity from spoken and written language samples (cf. Lu 2008; Roark, Mitchell and Hollingshead 2007; Wilde 2009).

Despite the persistent and heightened interest in the topic, complexity remains largely a notion that is not clearly theoretically defined and, even more importantly, one that lacks formal specifications (Juola 2008; Miestamo 2008). This is partly the consequence of its intuitive transparency, but also due to the multiplicity of theoretical frameworks which recently have been brought in to elucidate the topic. It is not surprising, then, that there is no single agreed-upon measure of either lexical or syntactic complexity.

Various alternative grammatical theories and models of human information processing make distinct predictions about the relative processing complexity of different syntactical constructions (Frazier 1989). One set of complexity measures reflects the assumptions that counting the constituent units of the input linguistic string (variably identified as morphemes, words and clauses) would give the researcher an

indication of the relative computational demands required to process the entire string. On the other hand, the motivation behind metrics based on developmental data (Lee 1974; Rosenberg and Abbeduto 1987) is that the difficulty of processing disparate syntactic structures determines the order of their acquisition, as documented on the basis of observational data. A third set of metrics, designed to closely reflect predictions derived from theoretical grammars, attempts to quantify the processing load associated with assigning a theory-grounded syntactic representation of input sentences, such as *X-bar syntax* (Frazier 1985), or to extract specific measures of dependency distance by using algorithms specified within *dependency locality theory* (Gibson 1998).

The nature of the processing mechanism responsible for the syntactic representation and the manner in which it interacts with the other subsystems of the language processing architecture is still a matter of debate. Under one approach to linguistic complexity, it is assumed that syntactic information is processed before other information in a specialised and informationally encapsulated module whose output "defines the task presented to the general cognitive processing system and places limitations on its solutions" (Frazier 1989: 200) (for a more recent account of modularity see Barrett and Kurzban 2006; Carruthers 2006). The constraint-based approaches, on the other hand, assume that a number of lexical, semantic and pragmatic factors guide the parsing process (e.g. Just and Carpenter 1992; MacDonald, Pearlmutter and Seidenberg 1994; Altmann 1998; Gibson 1998).

Researchers from either side of the debate generally agree that complexity effects in comprehension and production are due to WM limitations (e.g. Daneman and Carpenter 1980; Marcus 1980; Just and Carpenter 1992; Gibson 1998; Caplan and Waters 1999; Stowe, Withaar, Wijers, Broere and Paans 2002; Waters, Caplan and Yampolsky 2003). It remains an open question precisely how resources constrain sentence comprehension given the primary distinction between structural integration costs (associated with the attachment of currently processed words into the structure(s) computed so far) and storage for the partially processed structures (Gibson 1998).

Some researchers argue that all verbally-mediated tasks tap into one general pool of computational resources (Just and Carpenter 1992), and that variations in the capacity of linguistic working memory are the primary cause of individual differences in language processing. An alternative model proposes a specialisation in the verbal working memory system between (at least) two pools of resources (Caplan and Waters 1999). One resource system is dedicated to interpretative processes, i.e. automatic, first-pass language processes executed in a series of specialised modules. The other WM resources are utilised for conscious controlled processing of the propositional content of the sentence (referred to as post-interpretative processes), which is used to accomplish tasks such as decision-making, planning actions and long-term memory storage. The substantive evidence adduced in support of this model comes from findings, demonstrating the lack of certain statistical interactions between syntactic complexity and external processing load (Waters, Caplan and Yampolsky 2003).

Neither approach, however, appears to have a clear position regarding the role of experience and skill in language comprehension. Expertise research has challenged the resource-based accounts of processing. Ericsson and Kintsch (1995) have rejected the emphasis on working memory size in accounting for individual differences, arguing that perceived WM limitations can be circumvented through the acquisition of task-specific skills and knowledge following extended practice. Some of the empirical evidence for this hypothesis comes from text comprehension research which demonstrates that skilled comprehenders possess superior knowledge and strategies, which allowed them to construct more elaborate and better integrated representations of texts. Similarly, experts from a variety of cognitive domains are able to cope with the heavy demands on WM resources by relying on a combination of acquired skills, vast knowledge and task-specific processing strategies. The evidence in this respect points to expert's ability to strategically allocate attention to aspects of the task and its content, which are critical for its successful completion, to organise and encode the input in a manner anticipating future processing demands, and consequently gain a reliable and speedy access to their superior declarative and procedural knowledge on-line (Ericsson and Kintsch 1995; Ericsson and Lehmann 1996; for similar findings regarding expert interpreters, see Ivanova 1999). In this framework, individual differences in processing linguistic complexity are interpreted as differences in acquired skills as a function of experience.

Interpreting research has traditionally favoured capacity-constraint approaches to WM. Just and Carprenter's model informs the most extensive study to date of WM involvement in SI employing the paradigm of expert-novice differences (Liu 2001). Gile's Efforts Model of interpreting (Gile 1995) poses a single unitary limited-capacity resource pool, similar to the WM model of Just and Carpenter. Processing costs are specified in terms of four general distinct operations, called "efforts": comprehension, production, memory and coordination (responsible for resource allocation). Ineffective allocation of resources can result in "failure sequences" – instances when, for example, current production problems deplete resources necessary for the comprehension of the incoming parts of the message.

Psycholinguistic research on complexity is also grounded in approaches emphasising the limited nature of our cognitive resources. It has led to the formulation of a number of complexity metrics, devised to predict the difficulty of processing a range of syntactic constructions, and tested for usability in assessing linguistic development in first and second language acquisition and in diagnosing cognitive impairment. The eleven most widely used metrics have been evaluated by Cheung and Kemper (1992) for their discriminative adequacy in assessing age-related cognitive decline in sentence production and in predicting sentence comprehensibility and recall. The research demonstrated that eight of the metrics were highly correlated and can provide a reliable account for individual differences in linguistic complexity, viz. mean clauses per utterance (MCU), developmental sentence scoring (DSS), developmental level (D-Level), directional complexity (D-complexity), two variants of Yngve's depth metrics (reflecting the effect, "the depth", of embedding, esp. left-branching

structures) (Yngve 1960) and two variants of Frazier's node count (also based on the phrase marker of a given construction, but reflecting the number of non-terminal nodes that need to be introduced in order to integrate a word into the given parse tree). Cheung and Kemper specifically singled out *amount of embedding* and *type of embedding* as "adequate indexes of linguistic complexity" (1992: 74). Based on their conclusion, the two complexity metrics were selected for investigation in the present research.

Our experimental study was designed to ascertain whether and how aspects of syntactic complexity determine SI performance and to assess the effects of experience in dealing with processing difficulty as a function of linguistic complexity. Exploring the effect of complexity on SI offers a potentially propitious approach to the question of cognitive resource management and the specific nature of language processing underlying SI as an independent, albeit language-based task.

3. Experiment

3.1 Design

The research is designed as a quasi-experimental two-dimensional study (cf. Campbell and Stanley 1963), with the two dimensions being expertise and text. The expertise dimension involves comparing expert and novice interpreters. The text dimension, on the other hand, is a comparison of two texts (see below), very similar in most ways (i.e. speaker, content, presentation and linguistic features), but which differ in one specific way – their degree of information density.

3.1.1 *Subjects*

Two populations were sampled separately for the experiment – experts and novices, with the resulting two groups being homogeneous within the group, but well separated between the groups. The experts were identified using a set of factors such as accreditation, membership of professional organisation and peer-assessment. The novices were recruited among the students of a 2-year MA course in interpreting and had had three months of SI training, following a 3-year English major. Relevant aspects of the experimental population are summarised in Table 1.

Table 1. Relevant features of the experimental population

Profile features	**Group**	
	Experts	**Novices**
No. subjects (female)	8 (4)	8 (4)
Mean age	39	23
Average interpreting experience	9 years	3 months
Average years of active use of English	27	10

All subjects were fluent late bilinguals with Bulgarian as their mother-tongue and English as a B language, assessed at level C2 on the European Proficiency Scale. All subjects were using English daily for academic or work purposes. Both experts and novices were offered a participation fee.

3.1.2 *Procedure*

An experimental sound-proof booth was set up at a phonetics laboratory where each subject worked individually in the presence of the researcher. The external audio and recording equipment was manipulated by a technician; the subject had control over input volume and the external microphone.

Upon their consent to participate in the experiment, the subjects were presented with an interpreting brief detailing the topic of the speeches and the venue, also available at the start of the experiment. The experiment started with a three-minute warm-up text, which was the opening address taken from the same conference as the experimental texts and was produced by the same speaker (this was not subjected to further analysis). This gave the interpreters a chance to familiarise themselves with the speaker's idiolect, the topic of the conference and of the particular speech, all aspects of the conference macrotext which facilitate discourse processing (Ivanova 1993).

The texts were aurally presented from a recording in the subjects' B language. The warm-up text was consecutively followed by *Text1* and *Text2*. The material was not rotated for two specific reasons: firstly, because the number of subjects was small; and second, since one of the goals of the experiment was to study the effect of information density, we have determined (see below) that *Text2* is more redundant than *Text1,* so that rotating the materials could have potentially misbalanced the design. Interpreting was then followed by a debriefing interview, during which the subjects answered orally open-ended questions eliciting their individual assessment of their performance and addressing issues of ecological validity.

3.1.3 *Materials*

Two in situ conference audio recordings were selected for the experiment on the basis of the following characteristics:

- *Text type*: both are opening reports, therefore no detailed knowledge about the conference macrotext was necessary. Consequently, they could be presented in an experiment as relatively independent and self-sufficient discourses. Based on our estimate and responses of the subjects during the debriefing interviews, they can be considered typical of political conferences.
- *Level of scriptedness*: although both texts were planned and rehearsed, they were not fully scripted and exhibit a fair degree of orality.
- *Speaker:* the texts were produced by the same speaker (native British English) and delivered at two consecutive, but different conference events.
- *Topic:* the texts address a topical issue dealing with the integration of Eastern European states into the EU.

– *Specialised vocabulary:* no such vocabulary was found in either text; a list with all institutional, geographical and personal names appearing in both texts was presented to the subjects prior to interpreting.

A battery of lexical, syntactic and discourse parameters was manually and automatically derived to ascertain that the texts were linguistically comparable – a conclusion which is not surprising, given the fact that they were produced by the same speaker. Measures for the key linguistic and discourse characteristics of both texts are provided in Table 2.

Table 2. Linguistic profiles of the experimental texts (key to abbreviations: DL = Developmental Level; SDRS = Segmented Discourse Representation Structure; IP = Interactive Plane; * automatically derived values using Coh-Metrix, for detailed explanation of the metrics cf. Graesser, McNamara, Louwerse and Cai 2004)

	Text1	**Text2**
Lexical parameters:		
Total words*	572	648
Average syllables per word*	1.56	1.62
Mean lexical density (content words per segment)	18.6	24.6
Word types	192	193
Type: token ratio for content words*	0.71	0.62
Syntactic parameters:		
Mean DL	4.1	4.8
Number of segments*	21	17
Clauses*	62	68
Main	23	19
Subordinate	30	38
Mean clause density (clauses per segment)	3	4
Discourse parameters:		
No of SDRS	27	21
SDRS density (words per SDRS)	21.4	32.3
IP items	25	32
Words in IP structures (relative to total)	0.13	0.11

Finally, both texts scored 12 on the Flesh-Kincaid grade index of readability (computed with the help of Coh-Metrix, cf. Graesser, McNamara, Louwerse and Cai 2004), indexing the linguistic complexity of the text as a function of its syntactic properties. Despite the demonstrated linguistic similarity, *Text1* and *Text2* differed in one important aspect – redundancy.

3.1.4 *Redundancy: The text dimension in the study*

Substantiating the claim that both texts differ with respect to redundancy is not trivial, as there exists to date neither accepted procedure nor measures to estimate textual

informativity, even though the concept is frequently used in both discourse studies (cf. de Beaugrande and Dressler 1981) and interpreting research (Chernov 2004).

In the present study the notion of redundancy was operationalised in a multidimensional analysis of cohesion (interpreted as co-referentiality and argument overlap), new information units (discourse structures and new referents) and identification of pleonastic and parasitic discourse segments. As part of the effort to ensure the reliability of the comparison, the experimental texts were submitted for blind analysis to three linguists who were asked to identify new noun concepts (cf. Kintsch and van Dijk 1978) and to bracket information which could be justifiably removed from the texts without compromising the message content. The number of bracketed word tokens is calculated as a percentage of the total number of lexical tokens (the results are summarised in Table 3).

Table 3. Judges' estimates of redundancy for *Text1* and *Text2*

Judge	**New noun concepts**		**Redundant information items (word tokens as a percentage of all tokens)**	
	Text1	**Text2**	**Text1**	**Text2**
Judge1	66	46	20%	32%
Judge2	40	19	29%	42%
Judge3	54	29	22%	34%

The judges' analyses consistently demonstrate a higher redundancy for *Text2*. This is further collaborated by the higher incidence of discourse structures in *Text1* (see Table 2), indicating that it has a more complex and dense macro level structure and could consequently be expected to present a greater challenge for the high-level discourse processes. Additional metrics automatically computed with Coh-Metrix provided further support to the judges' analyses. Two measure of referential cohesion (Kintsch and van Dijk 1978) – argument and stem overlap – showed considerably greater overlapping of arguments, stems and content words for *Text2*. Latent Semantic Analysis across paragraphs (0.373 vs. 0.628 for *Text1* and *Text2* respectively) indicated greater semantic similarity among the content of *Text2* paragraphs. All these measures consistently demonstrate that *Text2* has a higher level of information redundancy, defined as "a large degree of interdependence among the successive units" (Miller 1963: 103).

Introducing the text dimension of the present study offers the opportunity to examine changes in processing strategies relative to manipulation of processing load and to access special features of expertise that "may only appear clearly with more difficult material or at faster rates of presentation" (Dillinger 1994: 184).

3.1.5 *Analysis of linguistic complexity*

Four complexity metrics were used as independent variables. Two measures of complexity were selected for the present study based on Cheung and Kemper's (1992)

psychometric evaluation discussed in Section 2. These were manually extracted by a linguist, but the structural annotation for the analysis of developmental level (arguably a critical point of analysis) was vetted by a syntactician. Linguistic complexity was also captured by two pairs of oppositions, contrasting sentences with different syntactic properties. The contributing syntactic factors were identified in previous SI research. A brief description of the metrics is given below.

Amount of embedding. This measure is based on mean clauses per utterance (MCU) proposed by Kemper and colleagues (Kemper, Kynette, Rach, O'Brien and Sprott 1989). Underlying this metric is the assumption that multiple syntactic embeddings contribute to processing difficulty and therefore adversely affect language comprehension and production. The measure was obtained by counting all clauses (both main and subordinate) per sentence, from which the sentences were classified as falling within three categories: low density (up to three clauses per segment, an intonationally, syntactically and semantically completed unit); middle density (between four and six clauses per segment); and high density (more than seven clauses).

Developmental Level (DL). The DL scale for assessing syntactic complexity was initially proposed by Rosenberg and Abbeduto (1987) using the acquisition order of syntax structures. Covington, He, Brown, Naçi and Brown (2006) expanded the scale to include a wider range of sentences along with minor revisions reflecting recent developmental evidence. Table 4 presents the updated scale, also used in this experiment, together with an inventory of the corresponding syntactic constructions relative to their position of embedding (subject or object place).

Table 4. Specification of the variable Developmental Level (DL) in terms of syntactic structures

Level	Description of syntactic constructions
0	simple, one-clause sentences, incl. questions, sentences with auxiliaries and semi-auxiliaries, simple elliptical sentences;
1	complex sentences with embedded infinitival components or -ing complement with the same subject as main clause;
2	complex sentences with wh-predicate complements, conjoined clauses with a coordinating conjunction, conjoint verbal, adjectival or adverbial constructions, and compound subjects;
3	complex sentences with relative clauses modifying the object noun phrase, nominalisation in object position, or with predicate noun phrase complements, subject extraposition;
4	complex sentences with gerundive complements or comparative constructions;
5	complex sentences with relative clauses modifying the subject noun phrase, subject noun phrase complements and subject nominalisations;
6	complex sentences with subordinate clauses;
7	multiple levels of embedding and subordination in a single sentence.

Clause type. This variable uses a relatively simple distinction between main and subordinate clauses and it designates the syntactic prominence of a clause. Main clauses are assumed to be easier to process (Frazier 1985; Dillinger 1994) because they are syntactically foregrounded, causing them to be more successfully interpreted.

Subordination type. The fourth syntactic variable introduces a distinction between two subordinate clauses, complement and relative. This contrast is not captured by the other metrics, although predictions have been made elsewhere concerning their relative processing difficulty. Thus, Frazier and Clifton (1996) argue that the different attachment procedures for complement and relative clause impose different processing demands on the parser. Under a generative account of grammar, parsing of complement clauses merely requires addition of a CP node as a sister to V. The attachment of a relative clause requires, on the other hand, "a new N' node to be postulated between the two already connected nodes of the phrase marker" (1996: 71). It is associated with a higher number of potential ambiguities, the resolution of which is contingent upon increased computational resources (Frazier and Clifton 1996).

3.2 Data manipulation

32 recordings were collected for both texts and transcribed using the Verbmobil repertory for spontaneous speech data (Burger 1997). The propositional analysis of the SL texts was integrated into a database where every proposition was coded for the properties of the segments in which they occurred. Performance data for every subject was appended as a new field to the database. To analyse performance accuracy (PA), the dependent variable in the study, four response categories (RCs) were used: strict correspondence, paraphrase, omission, and error. In the subsequent analysis the first two RCs were subsumed under the superordinate category of successful responses, while the last two were subsumed under unsuccessful. The reliability of PA coding was assessed by submitting 25% randomly selected protocols to five raters and using an algorithm described in Hild (2007). This yielded a high interrater agreement score – 90% correspondence between researcher's and raters' codings. Frequency data on RCs in the form of percentage was then analysed by means of a series of independent z-tests.

3.3 Results

The effect of the syntactic complexity on SI performance was first assessed independently of expertise and text (non-contrast condition). The results for the two complexity metrics, indexing amount and type of embedding, suggested a strong relationship between linguistic complexity and SI performance, similar to the one established by findings from monolingual research (Cheung and Kemper 1992). Predictably, increased complexity led to a systematic degradation of performance as a function of increased processing demands. For DL a strong correlation was established between PA

and DL (the Pearson correlation coefficient for a one-tailed/directional test $r = 0.75$, $p = 0.025$), with 58% of PA variation attributable to DL. When amount of embedding was assessed, significant differences were established for segments with high density vs. low ($z = 2.67$, $p < 0.01$) and middle density ($z = 3.23$, $p < 0.01$). No statistically significant results were obtained for subordination and clause type in the non-contrast condition.

However, for all metrics the results showed a strong effect of expertise and text. These will be discussed is separate sections below.

VARIABLE: DEVELOPMENTAL LEVEL (DL)
Values: [DL = 0:7]

The results reported for the non-contrast condition were replicated in the analysis of novices' performance on both texts where the same negative association of PA and DL was established (correlation significant at 2.5% level), with 60 % of PA variance attributable to DL (Figure 1).

A similar significant correlation was established for experts' performance on *Text2* ($r = 0.9$, $p < 0.01$). As Figure 1 shows, the data points align in an almost perfect descending line, indicating a decrease in performance accuracy from almost flawless renditions of segments with DL0 to a drop of 23% for those with DL7. However, for the condition of increased information load presented by *Text1* the predicted correlation was not obtained: linguistic complexity as indexed by type of embedding accounted for only 26% of the variance in interpreting accuracy.

VARIABLE: AMOUNT OF EMBEDDING
Values: [High/Middle/Low]

The z-tests on the novices' results show a large difference in the number of successful responses given to propositions in low and middle density vs. high density segments (for low-high density contrast $z = 2.67$; for middle-high density contrast $z = 3.61$; for both scores $p < .01$). This indicates that novices process high clause density segments in a different way from syntactically simpler segments, which contain up to six clauses. The findings were replicated for both texts. Separate analysis of sub-RC showed a significantly higher number of omissions for the condition with increased information load.

On statistical analysis, the experts' data yield only one moderately significant result when low and high density segments were compared ($z = 2.01$, $p < 0.05$). Similar findings were obtained for *Text2*, but not for *Text1* where none of the investigated contrasts reached significance.

Overall, the variable has a much stronger and consistent effect on the performance of the novices, whose PA deteriorates dramatically as a function of the number of embedded clauses per segment (cf. Figure 2).

VARIABLE: CLAUSE TYPE
Values: [Main/Subordinate Clause]

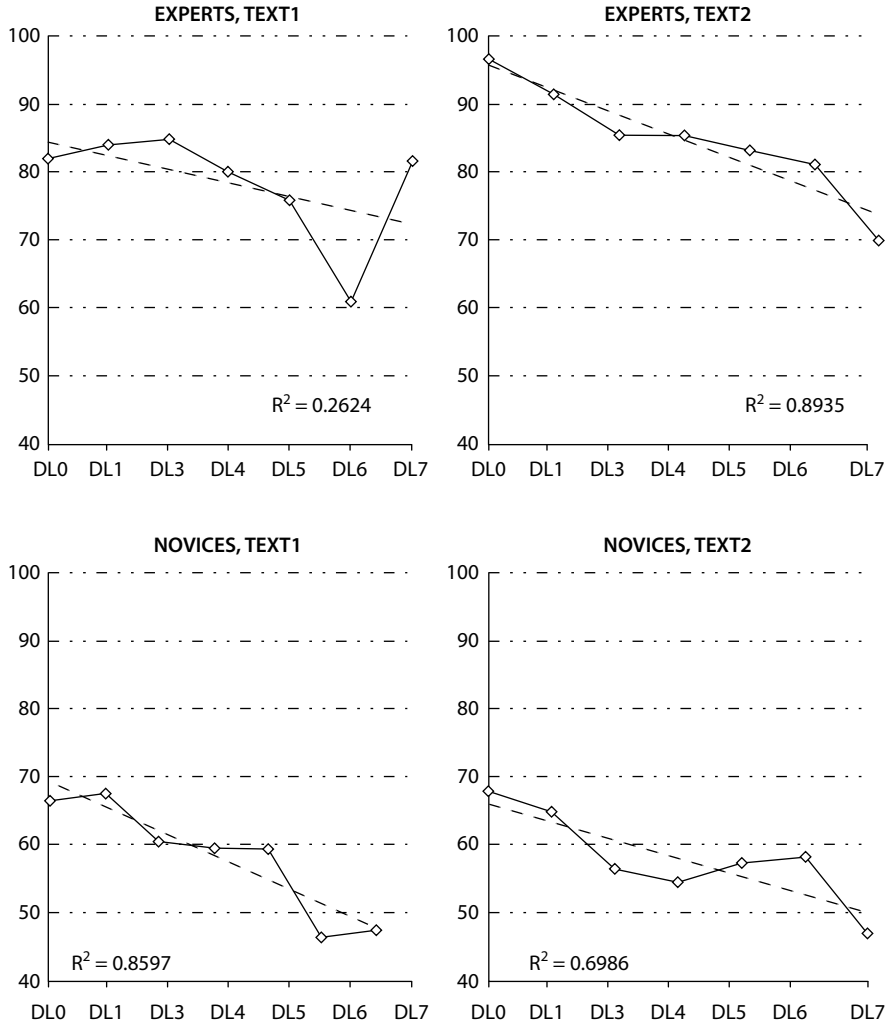


Figure 1. Trend analysis of the effect of DL and text on expert and novice performance

The study of the variable was based on the prediction that syntactically embedded clauses are more difficult to process, and consequently they would be interpreted less accurately. When the effect of the variable was assessed in relation to experience and text, the following pattern emerged (Figure 3).

As Figure 3 shows, the variable had a weak effect on experts' performance (no statistically significant differences emerged from the analyses) but a consistent effect on that of the novices. The z-tests carried out on the novices' consolidated results showed that propositions in subordinate clauses have a significantly lower probability of being correctly interpreted ($z = 2.79$, $p < 0.01$) and are omitted more frequently ($z =$

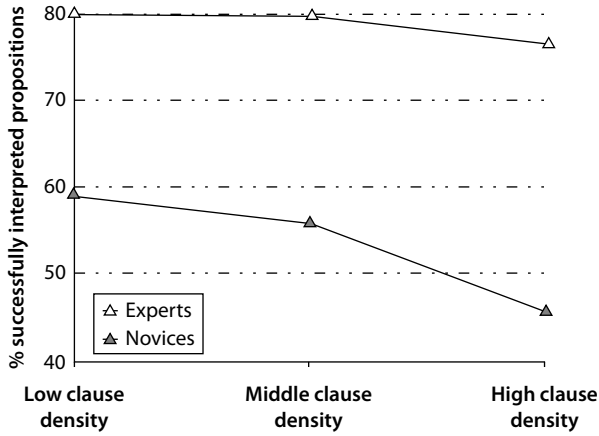


Figure 2. Joint effects on interpreting of amount of embedding, experience and text

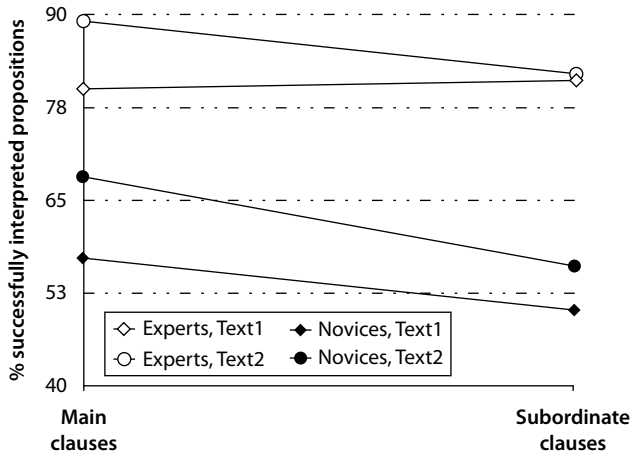


Figure 3. Joint effects of Clause Type, experience and text on interpreting

2.11, $p < 0.05$) than propositions in main clauses. This is true for *Text2* expressed in the increased proportion of successful ($z = 2.32$, $p < 0.05$) and strictly corresponding responses ($z = 2.48$, $p < 0.05$) for main clauses. The results from *Text1* did not reach significance level, but they followed the main trend; i.e. the novices gave quantitatively more successful responses to propositions in main clauses. On the whole, interpreting accuracy for *Text1* was decreased, suggesting novices' performance was at ceiling in the condition of increased load.

VARIABLE: SUBORDINATION TYPE
Values: [Relative/Complement Clause]

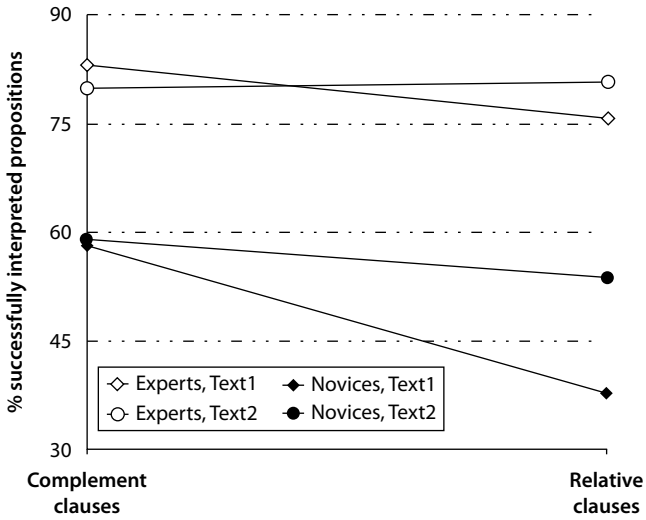


Figure 4. Joint effects of Subordination Type, experience and text on interpreting

The effect of subordination on interpreting was further explored by investigating the interaction of two types of subordinate clauses (relative and complement clauses) with interpreting performance (Figure 4). The z-tests on the experts' results for both texts showed no statistically significant differences in the number of successful responses given to information in either type of subordinate constructions. This is true for each text. Therefore, this syntactic distinction appears to have no effect on experts' processing – complement and relative clauses (R-clause) are being equally well processed, regardless of their syntactic status.

In contrast, the novices' performance is sensitive to this syntactic distinction. When the effect of the distinction was examined in the consolidated condition, the z results showed a significant difference in the number of successful responses given by novices to propositions in complement in contrast to relative clauses ($z = 2.6$, $p < 0.01$). The analysis of R-clauses suggests that this difference in performance is due to the higher omission rate for propositions in relative clauses ($z = 2.2$, $p < 0.05$). The same was true of the novices' results on *Text1*. In the no-load condition the tendency to interpret relative clauses less accurately was preserved, but the tests did not yield any statistically significant results.

Furthermore, a post-hoc linguistic analysis of the novices' transcripts was undertaken to diagnose the problems related to RC processing. The most frequent error type, which occurred at the attachment sites between RC and its head, involved violation of the predication rule which specifies the agreement requirements for wh-complementisers (Haegeman 1994). In English agreement is not explicitly marked in contrast to the TL – Bulgarian, where head and wh-complementisers carry overt morphological markers of agreement with respect to gender, person and number.

Inferring the implicit grammatical information on-line may take a considerable amount of time in part because the required information is provided at a later time in the sentence. The processing outcome becomes contingent upon an interpreter's ability to use a combination of semantic and syntactic cues in conjunction with an extended décalage, thus multiplying the processing costs for both integration and storage. Consider the following example of object R-clauses (the relative pronoun extracted from the object position of the embedded clause):

SL [$_{NP}$ two *communications*$_1$ [$_{CP}$ ***which*$_1$** [$_{IP}$ the Commission [VP made$_{t1}$ to the council]]]] … *(–singular) (–singular)*

TL [$_{NP}$ dve *komunikazii,* [$_{CP}$ ***koeto*** [$_{IP}$ [$_{VP}$ napravi$_{t1}$] Komicijata]]] … [two *communications (–singular)*, ***which*** *(+singular, +neuter)* did the Commission]

The grammatical features of the TL relative pronoun in the example make it impossible to identify its dominating head and establish co-referential links to any of the preceding NPs. Following Gibson (1998), we propose that the complexity effects in the example above can be explained in terms of the increased memory cost of processing incomplete syntactic dependencies: the embedding of the subject "Commission" between the relative pronoun and its role-assigning verb and integration of the R-clause with the governing head "communications". The simplified syntactic structure of the TL R-clause and the error in thematic structure assignment for the R-clause verb suggest that the novice failed to process all linguistic cues ensuring correct integration of the R-clause prior to the commencement of translation. While it might be argued that object R-clauses count as one of the more challenging syntactic structures in terms of their WM demands, it should be pointed that attachment errors of the type discussed above were also observed for subject R-clauses. Generally considered simpler because they are locally integrated, subject R-clauses pose no processing challenges either for healthy or impaired individuals (Gibson 1998). We can therefore speculate that the degraded performance on R-clauses is not so much a consequence of capacity constraints as a result of task-induced processing limitations characteristic of unskilled performance, evidence for which has been collected for a variety of cognitive domains (cf. Ericsson and Kintsch 1995).

4. Discussion

The motivation behind this study was to approach in a more proximal manner the nature of skilled interpreting by exploring the effects of cognitive load on SI processing. In looking for determinants of processing difficulty, the study turned to research on linguistic complexity and psychometric evaluation of complexity metrics. The selection of measures for assessment in SI was guided by their demonstrated utility in accounting for variations in comprehension and recall, on the basis of earlier

psycholinguistic research. Four syntactic factors were thus identified: amount and type of embedding, and clause and subordination type. The first two appear to determine how accurately SL sentences were processed, although the significant relationship established between linguistic complexity and SI performance was modified by expertise and information load.

The experiment established that all syntactic factors investigated in the study consistently affect novices' performance, suggesting the importance of syntactic information in inexperienced interpreting. Novices' processing showed the same pattern of interaction between complexity and performance established for monolingual processing (Cheung and Kemper 1992). Increasing further processing difficulty by presenting the novices with an informationally denser text significantly decreased accuracy, but did not alter the nature of these interactions. These findings replicated results from studies which examine the effect of concurrent memory load on syntactic processing (Caplan and Waters 1999) as an evidence for a specialised WM system utilised in syntactic analysis and provided experimental evidence for the observations (Gran 1989) that novices employ form-dependent strategies.

The findings, however, were not replicated for expert SI processing, where the effects of linguistic complexity varied with information load and type of measurement. The two bona fide metrics (*amount* and *type of embedding*) provided an adequate account of variations in performance accuracy for experts under the condition of decreased informativity. However, the relationship between complexity and performance breaks down when processing load was increased by presenting an informationally denser text. This, together with the absence of effects of subordination and clause on PA, suggests that expert processing is less vulnerable to linguistic complexity and that factors other than the syntactic properties of the ST become more important determinants of skilled SI performance. Our findings thus concur with Setton's conclusion regarding the attenuated effect of syntactic factors drawn from corpus analysis of output produced by experienced professional interpreters: "syntactic structure … does not itself constitute an obstacle to SI" (Setton 1999: 270).

One possible account for the findings is in terms of experts' increased reliance on post-interpretative processes in the sense of Caplan and Waters (1999) – processes that deal with the deployment of results from language processing to achieve SI-specific goals, such as minimising information loss and enhancing TT coherence. An alternative explanation suggests that skilled performance is modulated by more sophisticated and complex expertise-specific cognitive strategies. More speculatively, these will critically rely on the experts' ability to encode information selectively from the ST in anticipation of future retrieval demands and to efficiently deploy macrostrategies to enable the formation of macrostructures (Kintsch 1998), a conclusion supported by the obtained overall effect of text on processing. The effect of redundancy and text furthermore demonstrates experts' flexibility in deploying comprehension skills for optimising SI performance.

Our results point in the direction of systematic differences in the comprehension of complex syntactic structures between expert and novice interpreters, differences

modulated not by innate WM size, but by the availability of acquired skills and knowledge. Furthermore, the use of empirically verified complexity metrics gave us the opportunity to test several predictions derived from research on monolingual language processing (normal text comprehension) in the SI conditions. Overall, they were not supported by expert data, but were borne out by the results of the novices. The divergence of outcomes further demonstrates that text processing in expert SI is qualitatively different from normal text comprehension, while the unskilled SI still continued to rely on general-purpose strategies and skills characteristic of text comprehension.

Finally, the established reliability of complexity metrics in predicting novice performance has potential implications for training. Complexity metrics can be used in the selection and grading of appropriate training materials and to provide an objective approach to tracing the progress of the would-be interpreters. These possibilities are particularly attractive given the recent advances in development of computational tools (cf. Coh-Metrix, Computerized Profiling, www.computerized-profiling.org) used to automatically measure syntactic complexity.

References

Altmann, G. T. M. 1998. "Ambiguity in sentence processing." *Trends in Cognitive Sciences* 2 (4): 146–152.

Barik, H. 1975. "Simultaneous interpretation: Qualitative and linguistic data." *Language and Speech* 18 (3): 272–297.

Barrett, H. C. and Kurzban, R. 2006. "Modularity in cognition: framing the debate." *Psychological Review* 113 (3): 628–647.

Burger, S. 1997. *Transliteration of Spontaneous Speech Sata: Manual of Transliteration Conventions.* University of Tübingen: Verbmobil Phase II.

Campbell, D. T. and Stanley, J. C. 1963. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally & Co.

Caplan, D. and Waters, G. 1999. "Verbal working memory and sentence comprehension." *Behavioral and Brain Sciences* 22 (1): 77–94, discussion 95–126.

Carruthers, P. 2006. *The Architecture of the Mind.* Oxford: Oxford University Press.

Chernov, G. V. 2004. *Inference and Anticipation in Simultaneous Interpreting*. R. Setton and A. Hild (eds). Amsterdam/Philadelphia: John Benjamins.

Cheung, H. and Kemper, S. 1992. "Competing complexity metrics and adults' production of complex sentences." *Applied Psycholinguistics* 13: 53–76.

Covington, M. A., He, C., Brown, C., Naçi, L. and Brown, J. 2006. "How complex is that sentence? A proposed revision of the Rosenberg and Abbeduto D-Level Scale." *ReVision.* Washington, D.C. http://www.ai.uga.edu/caspr/2006-01-Covington.pdf. (Accessed May 10, 2010.)

Daneman, M. and Carpenter, P. 1980. "Individual differences in working memory and reading." *Journal of Verbal Learning and Behavior* 19: 450–466.

de Beaugrande, R. and Dressler, W. 1981. *Introduction to Text-Linguistics*. London: Longman.

Dillinger, M. 1994. "Comprehension during interpreting: What do interpreters know that bilinguals don't?" In *Bridging the Gap: Empirical Studies in Simultaneous Interpretation*, S. Lambert and B. Moser-Mercer (eds), 155–189. Amsterdam/Philadelphia: John Benjamins.

Ericsson, K. A. and Kintsch, W. 1995. "Long-term working memory." *Psychological Review* 102 (2): 211–45.

Ericsson, K. A. and Lehmann, A. 1996. "Expert and exceptional performance: Evidence of maximal adaptations to task constraints." *Annual Review of Psychology* 47: 273–305.

Frazier, L. 1985. "Syntactic complexity." In *Natural Language Processing: Psychological, Computational and Theoretical Perspectives,* D. Dowty, L. Karttunen and Z. Arnold (eds), 129–189. Cambridge: Cambridge University Press.

———. 1989. "The study of linguistic complexity." In *Linguistic complexity and text comprehension: Readability issues reconsidered*, A. Davison and G. Green (eds), 193–221. Hillsdale, N.J.: Erlbaum.

Frazier, L. and Clifton, C. 1996. *Construal*. Cambridge, Mass.: MIT Press.

Gernsbacherr, M. A. and Shlesinger, M. 1997. "The proposed role of suppression in simultaneous interpretation." *Interpreting* 2 (1/2): 119ff.

Gerver, D. 1971. "Aspects of simultaneous interpretation and human information processing." (Doctoral dissertation, Oxford University, UK.)

———. 1974. "Simultaneous listening and speaking and retention of prose." *Quarterly Journal of Experimental Psychology* 26: 337–341.

Gibson, E. 1998. "Linguistic complexity: Locality of syntactic dependencies." *Cognition* 68 (1): 1–76.

Gile, D. 1995. *Basic concepts and models for interpreter and translator training.* Amsterdam/Philadelphia: John Benjamins.

Graesser, A. C., McNamara, D. S., Louwerse, M. M. and Cai, Z. 2004. "Coh-metrix: analysis of text on cohesion and language." *Behavior Research Methods, Instruments & Computers: A Journal of the Psychonomic Society*, Inc 36 (2): 193–202.

Gran, L. 1989. "Interdisciplinary research on cerebral asymmetries: significance and prospects for the teaching of interpretation." In *The theoretical and practical aspects of teaching conference interpretation*, L. Gran and J. M. Dodds (eds), 93–100. Udine: Campanotto.

Haegeman, L. M. V. 1994. *Introduction to Government and Binding Theory*. 2nd edition. Oxford: Blackwell.

Hawkins, J. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Hild, A. 2007. "Establishing rigour in a between-method investigation of SI expertise." In *Doubts and Directions in Translation Studies*, Y. Gambrier, M. Shlesinger and R. Stolze (eds), 99–112. Amsterdam/Philadelphia: John Benjamins.

Ivanova, A. 1993. "Intertextuality in simultaneous interpreting: A case-study of two scenarios of the conference discourse." (Master's thesis, University of Sofia, Bulgaria.)

———. 1999. "Discourse processing during simultaneous interpreting: An expertise approach." (Doctoral dissertation, University of Cambridge, UK.)

Juola, P. 2008. "Assessing linguistic complexity." In *Language Complexity: Typology, Contact, Change*, M. Miestamo, K. Sinnemaeki and F. Karlsson (eds), 89–109. Amsterdam/Philadelphia: John Benjamins.

Just, M. A. and Carpenter, P. A. 1992. "A capacity theory of comprehension: individual differences in working memory." *Psychological Review* 99 (1): 122–149.

Kemper, S., Kynette, D., Rach, S., O'Brien, K. and Sprott, R. 1989. "Life-span changes to adults' language: Effects of memory and genre." *Applied Psycholinguistics* 10: 49–66.

Kintsch, W. 1998. *Comprehension: A Paradigm for Cognition*. Cambridge, UK: Cambridge University Press.

Kintsch, W. and van Dijk, T. A. 1978. "Towards a model of text comprehension and production." *Psychological Review* 85: 363–394.

Lee, L. 1974. *Developmental Sentence Analysis*. Evanston, Ill.: Northwestern University Press.

Liu, M. 2001. "Expertise in simultaneous interpreting: A working memory analysis." (Doctoral dissertation, University of Texas at Austin, Texas.)

Lu, X. 2008. "Automatic measurement of syntactic complexity using the revised developmental level scale." *Applied Psycho Linguistics*, 153–158.

Lyons, K., Kemper, S., LaBarge, E., Ferraro, F., Balota, D. and Storandt, M. 1994. "Oral language and Alzheimer's disease: A reduction in syntactic complexity." *Aging and Cognition* 1 (4): 271–281.

MacDonald, M. C., Pearlmutter, N. J. and Seidenberg, M. 1994. "Lexical nature of syntactic ambiguity resolution." *Psychological Review* 101 (4): 676–703.

Marcus, M. 1980. *A Theory of Syntactic Recognition for Natural Language.* Cambridge, Mass.: MIT Press.

Miestamo, M. 2008. "Grammatical complexity in a cross-linguistic perspective." In *Language Complexity: Typology, Contact, Change,* M. Miestamo, K. Sinnemaeki and F. Karlsson (eds), 23–43. Amsterdam/Philadelphia: John Benjamins.

Miller, G. A. 1963. *Language and Communication*. New York: McGraw-Hill.

Roark, B., Mitchell, M. and Hollingshead, K. 2007. "Syntactic complexity measures for detecting Mild Cognitive Impairment." *Computational Linguistics* 6: 1–8.

Rosenberg, S. and Abbeduto, L. 1987. "Indicators of linguistic competence in the peer group conversational behavior of mildly retarded adults." *Applied Psycholinguistics* 8: 19–32.

Sampson, G., Gil, D. and Trudgill, P. 2009. *Language complexity as an evolving variable.* Oxford: Oxford University Press.

Seeber, K. G. 2005. "Temporale Aspekte der Antizipation beim Simultandolmetschen komplexer SOV-Strukturen aus dem Deutschen." In *Empirical Research into Translation and Interpreting: Products and Processess* [Bulletin Suisse de linguistique appliqué 81], A. Künzli (ed.), 123–149.

Setton, R. 1998. "Meaning assembly in simultaneous interpretation." *Interpreting* 3: 163–199.

———. 1999. *Simultaneous Interpreting: A Cognitive-Pragmatic Analysis.* Amsterdam/Philadelphia: John Benjamins.

Stowe, L. A., Withaar, R. G., Wijers, A. A., Broere, C. A. J. and Paans, A. M. J. 2002. "Encoding and storage in working memory during sentence comprehension." In *The Lexical Basis of Sentence Processing: Formal, Computational and Experimental Issues*, P. Merlo and S. Stefenson (eds), 181–205. Amsterdam/Philadelphia: John Benjamins.

Uchiyama, H. 1991. "Problems caused by word order when interpreting/translating from English into Japanese: The effect of the use of inanimate subjects in English." *Meta* 36: 406–413.

Waters, G., Caplan, D. and Yampolsky, S. 2003. "On-line syntactic processing under concurrent memory load." *Psychonomic Bulletin & Review* 10 (1): 88–95.

Wilde, L. E. 2009. "A comparison of seven automated measures of syntactic complexity." (Master's thesis, Brigham Young University.)

Yngve, V. H. 1960. "A model and a hypothesis of language structure." *Proceedings of the American Philosophical Society* 104: 444–466.

Process and product in simultaneous interpreting

What they tell us about experience and expertise*

Elisabet Tiselius[#] and Gard B. Jenset[##]
[#]University of Bergen, Norway / [#]Stockholm University, Sweden
[##]Bergen University College, Norway

The expertise approach (Ericsson 2008) has been used to explore the competence of translators and interpreters since the mid-1990s, and is now a well established sub-field in translation and interpreting process research (Jääskeläinen 2010). In the area of interpreting, Ivanova (1999), Liu (2001) and others have explored the expertise approach. The studies reported in this article follow up on this work, but go one step further and investigate both process and product. The aim of the two studies was to explore the differences in performance between interpreters with shorter and longer experience (possible experts). Participants (n = 9) with no, short or long experience interpreted the same speech and performed retrospection immediately after. The first study, dedicated to process, used Ivanova's (1999) method for investigating the process. The second study, on product, let two groups, non-interpreters (n = 6) and interpreters (n = 6), rate the interpreting performances using Carroll's (1966) scales for intelligibility and informativeness. It was found that the degree of experience influences the processing strategies used by interpreters and the types of problems they report. Experience also has an impact on how the product of experienced interpreters and that of less experienced interpreters is rated, both when rated by interpreters and by non-interpreters.

Keywords: expertise, simultaneous interpreting, retrospection, process, monitoring, strategies, rating

---

* The authors would first like to thank the interpreters that participated in this study. The authors are also indebted to Birgitta Englund Dimitrova, Åse Johnsen, Barbara Blair, and Christer Johansson as well as the anonymous reviewers of this article for their insightful comments and corrections. Any faults or weaknesses that remain despite their close scrutiny are of course the responsibility of the authors alone.

1. Introduction

Expertise develops over time, as performers develop experience in their field. Ericsson (2000) outlines the expertise approach as follows:

> The theoretical framework of expert performance makes the fundamental claim that improvement in superior reproducible performance of adult experts doesn't happen automatically nor magically. Performance improvement can be linked to changes in cognitive mechanisms mediating how the brain and nervous system control performance and in the degree of adaptation of physiological systems of the body. The principal challenge to attaining expert level performance is to induce stable changes that allow the performance to be incrementally improved.
> (2000: 201–202)

While experience can easily be described in temporal terms, e.g. the number of years in the profession, hours of training or repetitions of the task, it is more delicate to describe or identify expertise. In their work on a theory of expertise Ericsson and Smith (1991) listed six criteria to be taken into consideration in order to study expert performance: long experience, consistent outstanding performance, access to expert knowledge when needed, deliberate practice, clear goals, and feedback. In competitive fields that involve ranking, such as tennis or chess, an examination of these criteria for the top performers will in all probability identify the experts in the field. Their performance can then be examined in order to identify the features of expert performance. However, it is more difficult to identify the possible experts in fields in which ranking is not available as a means to investigate their expertise. One method is to compare the work of experienced practitioners with that of novice practitioners, and map the former's performance and background. The differences (if any) between the two groups may constitute clues to outstanding practice in that field. On the basis of those differences, it may be possible to describe expertise and expert performance in the field. As Ericsson (1991: 8) puts it, "the expertise approach is an attempt to describe the critical performance under standardized conditions, to analyze it, and to identify the components of the performance that make it superior." In order to investigate the performance of interpreters by employing the expertise approach, it is necessary to consider both process and product: process, because these may alter with experience without necessarily showing in the product; and product, because the products of experienced practitioners most likely differ from those of less experienced practitioners. From these two angles it becomes possible to identify the components of superior performance, and consequently describe expertise in the relevant field, in this case simultaneous interpreting.

1.1 Aim and scope

The overall aim of the two studies presented in this paper is to explore differences in performance between inexperienced and experienced interpreters. This has certainly been done before (see Section 2.1 below), but the field is not yet completely mapped; proceeding step by step, it may be possible to identify the components of superior performance.

The overall question that the two studies sought to answer was: Can a difference in performance be established between three groups of interpreters with different levels of experience?

Previous research suggests that there is likely to be a difference in performance that depends on the interpreter's experience. The crucial question is: can this difference be identified and demonstrated in the context of an experimental design? There were two research questions:

First study: Is there a difference in the interpreting process (manifested through reported processing problems, instances of monitoring and strategies (see Appendix 1)) when carried out by groups of interpreters with different experience? Previous studies indicate that there is a difference in the interpreting process employed by experienced interpreters and inexperienced interpreters, see e.g. Ivanova (1999) and Vik-Tuovinen (2006). Years of practice and knowledge of the field are likely to make a difference. However, can this difference be demonstrated in an experiment that taps into the process? Is it possible to identify what the difference consists of?

Second study: Is there a difference in the ratings of the interpreting products produced by groups of interpreters with different levels of experience? Again, on the basis of previous research, such as the ones mentioned above, it is reasonable to hypothesize that there is a measurable difference between the groups investigated. The question remains: can this assumed difference be demonstrated in the context of an experiment?

2. Background

2.1 Expertise research in Interpreting Studies

An important focus of studies on conference interpreting has been on comparing inexperienced interpreters with experienced interpreters (e.g. Gerver 1971; Anderson 1979; Dillinger 1989). In the late 1990s Moser-Mercer (2000) took the concept of expertise from cognitive psychology into interpreting research, and since then several doctoral theses have been defended on that topic (see for instance Ivanova 1999 or Liu 2001). When exploring expertise, studies have, for instance, analyzed the interpreters' performances either by studying the process (e.g. Ivanova 1999) or the product (e.g. Liu 2001) quantitatively, or by analyzing the product and the process both qualitatively and quantitatively (Vik-Tuovinen 2006).

Ericsson (1991) includes three steps in his original expertise approach. The first is to capture the essence of superior performance under standardized laboratory conditions by identifying representative tasks. The second step is to provide a detailed analysis of the performance and the third is to account for the acquisition of the characteristics and cognitive structures and processes that have been found to mediate the superior performances of experts (1991: 12). These steps may sound straightforward, but represent a major challenge for research into interpreting. Interpreting consists of a multitude of tasks all performed simultaneously (cf. the Moser model 1978). Only by testing interpreters' performance under relatively standardized conditions is it possible to isolate and analyze the superior performance identified in step two. In Ericsson's words, "by careful analysis of the expert performance in real life, we try to identify recurrent activities that can be reproduced under controlled conditions" (1991: 14). The two studies reported here are parts of Ericsson's first and second step. We study the performance of all the interpreters under reasonably standardized conditions. We contrast the different performances, analyze them, and explore possible superiority.

Ivanova (1999: 182), whose method is described in greater detail in the next section, saw that experienced interpreters encounter fewer processing problems and have more elaborate problem representations. When it comes to problem solving and strategies they have a wide range of context-sensitive solutions at hand. Ivanova concluded that an important feature of interpreting expertise is its adaptive nature relying on the cognitive flexibility of the experts. Liu (2001: 93), who compared 13 professional interpreters (with training and at least 2 years of experience) with two student groups (n = 12 and n = 11), found that experienced interpreters had developed domain-specific skills (ability of selective encoding, better monitoring of output, more efficient allocation of working memory), which were in turn crucial in outperforming less experienced subjects. Vik-Tuovinen (2006) investigated seven experienced interpreters with one to fifteen years of experience and compared them to two groups of six and eight students respectively. She found that beginners focus on the source text and the linguistic expression, whereas professional interpreters focus more on social aspects and situational factors (Vik-Tuovinen 2006: 313).

Liu (2008), in a general overview of research on interpreting expertise, concluded that expert interpreters differ from novices in that they "have developed well-practiced strategies in each of the comprehension, translation and production processes." However, a crucial part of their expertise is not only the development of these strategies, but the skill to manage them during the task of simultaneous interpreting.

2.2 Studying the process

Ericsson states that "comparison of think-aloud verbalizations by experts and novices is the best-known method of assessing differences in the mediating processes as functions of the subjects level of expertise" (1991: 20). When investigating the process of interpreting there are not many methods at hand. A pure think-aloud method while

the subject is performing a task is impossible; the interpreter cannot both interpret and verbalize on the process at the same time. There are very limited possibilities to tap into the on-going process of interpreting; none if we want the interpreting subjects to verbalize their process. Immediate retrospection, however, is a possibility since the subjects recall and verbalize their memories of the process after the task. The retrospective interview is usually prompted by a cue; for interpreting this can be either the source text or the interpreter's own production (sometimes supported by the source text), as suggested by Ericsson and Simon (1996).

Both Ivanova (1999) and Vik-Tuovinen (2006) use retrospection in their studies of expertise in interpreting. The first study reported in this chapter closely follows the method used by Ivanova in her PhD thesis (1999). Her subjects (n = 16, including novices as well as possible experts) interpreted an English speech into Bulgarian; immediately after interpreting, the subjects performed a retrospective interview where they used only the source text as cue. Following a pilot study comparing the reliability of two types of cues (source text only and source text with a target text recording), Ivanova decided to use source text only since she found that use of the target text recording ran the risk of influencing subjects, who could deduce the process from the product rather than recalling the process in itself. For an examination of Ivanova's method and retrospection for both translation and interpreting, see Englund Dimitrova and Tiselius (2009).

2.3 Evaluating the product

Quality and how to asses it has been discussed by researchers such as Moser-Mercer (1996), Shlesinger (1997) and Gile (2003). Shlesinger's article referred to here sums up a discussion on quality in which many eminent members in the interpreting research community participated. There seems to be a general consensus that quality cannot be assessed in absolute terms: it has to be handled from many different angles, not only the communicative event, naturally, but also its fidelity to the information in the source speech and its value as a speech of its own.

Quality is a well researched topic in interpreting studies. Several investigations have been made in the field and many methods for evaluating the product have been developed. Bühler (1986) conducted the first larger-scale survey on interpreters' own perception of quality, while Kurz (2002: 323) undertook a similar investigation in different user groups. The result of these surveys was somewhat similar in that interpreting is good if it serves its purpose; an ideal interpreting is not an absolute value but depends on e.g. situationality and communicative context. The studies conducted by Bühler and Kurz are now being followed up by a doctoral thesis on quality and the role of the interpreter (Zwischenberger, Pöchhacker and Kurz 2008). Clearly, the quality in this case is investigated by means of audience response, with both interpreters and non-interpreters as audience. The more recent, regular and larger studies of quality from this perspective are done by the Interpreting Services at the European

Commission (SCIC) in its regular customer satisfaction survey (2007, 2010). Among the studies from the customers' perspective are also Ng (1992) and Vuorikoski (1998). In the study reported here on assessment of the product, both the interpreter's perspective and the user's perspective were taken into account.

Relating quality to the investigation of expertise, Ericsson (1991: 15) states that "although judges can reliably assess the superior quality of the product, it is difficult to analyze such products in order to identify the measurable aspects capturing the superior quality of the product." Hence, it is important to combine the assessment of quality with the investigation of the cognitive processes underlying the production of an interpreting of a certain quality.

In the second study presented here below, Carroll's scales (1966) were used as tools to assess the products of the translation (see below). Carroll's scales were originally designed to measure the quality of machine translation. In the present study they have been adapted to interpreting. There are two scales: one measures the *intelligibility* of the product on its own, while the other measures the *informativeness* of the product compared to the original.

For a detailed presentation and investigation of Carroll's scales, their adaptation to interpreting and their appropriateness for this type of assessment, see Tiselius (2009).

2.4 The studies

In this article we report on two studies, one focusing on the process of interpreting, and the other one focusing on the product. The same data was collected for the two studies: nine subjects with different levels of experience interpreted the same nine-minute-long speech. Immediately after interpreting, they carried out a retrospection task using a written transcription of the original speech as cue. The interpreting performances and the retrospections were recorded and transcribed. In the first study the interpreting performances were compared across the different levels of experience. In the second study the interpreting products were assessed by 12 different assessors, both interpreters and non-interpreters, and then the performance ratings were compared.

3. First study: Interpreting process

In the first study we investigated the differences in the interpreting process in the groups of interpreters with different levels of experience. We took the retrospections as a starting point in our analysis and classified the reported instances from the retrospections into Processing Problems, Monitoring and Strategies, following Ivanova's (1999) method (see Appendix 1). These instances were then compared across the three groups in order to investigate whether the pattern differed depending on the level of experience.

3.1 Material and method

3.1.1 *Input material*

The speech to be interpreted was held by Commissioner Byrne in the European Parliament in May 2002 and was approximately nine minutes long (00:09:35). For the purpose of the present study it was slightly adapted and rerecorded using another speaker. Some names and figures were added in order to increase the cognitive load, and the pace was adjusted to a slower speech rate in order to make it possible for all of the subjects to interpret. The original speech had an average speed of 141 words-per-minute and the speaker had a pronounced Irish accent. The re-recorded speech was delivered at an average speed of 119 words-per-minute[1] in Received Pronunciation (RP).

As Ivanova (1999) points out, background knowledge is important for an expert performance, enabling the expert to anticipate and to draw conclusions from previous experience. For this reason Ivanova chose a bona fide conference speech. This is also the reason why a speech from a real-life situation was chosen for the present study. The written transcript of the speech is provided in Appendix 2.

The speech was recorded on a CD in a sound proof studio, with a frequency range of about 20 000 Hz, which is an important criterion for good interpreting working conditions, as compared with the frequency range of a telephone line, which is about 4 000 Hz.

3.1.2 *The interpreting subjects*

The participants in this study (n = 9) had three different levels of interpreting experience; no, short and long. Those with short and long experience had both training and experience, while those with no experience did not have any interpreting experience. The no experience group was recruited among translation students beginning their second semester at the Institute for Interpretation and Translation Studies (TÖI) at Stockholm University. Students come to TÖI after at least one year of undergraduate language studies. They had seen the inside of an interpreting booth during an introductory lesson to conference interpreting, where they spent 15 minutes in the booth trying to interpret an easy text from English into Swedish. The interpreters with short experience were recruited among the former students of the European Masters in Conference Interpretation at TÖI. All of them had two years of experience at the time of the recording. The interpreters with long experience were recruited among the interpreters at the European Parliament and all had at least 25 years of experience (see Table 1).

The interpreters with long experience in this study clearly have credentials that would indicate some level of expertise: very long experience; accreditation at international institutions; and membership in the international conference interpreters'

1. There seems to be a limit to how fast interpreters can actually process an utterance. Seleskovitch (1965) suggested that 100–120 words per minute (wpm) is a comfortable speed, and 150–200 the upper limit. Gerver's study (1969, 1971) supported her estimate.

Table 1. Age and experience of the interpreters

Group	Age span	Years at university	Interpreting school	Years of experience
No experience	20–29	4	No	0
Short experience	30–49	4	Yes	2
Long experience	50–60	4	Yes	25+

organization, AIIC, where prospective members must be sponsored by full members to obtain membership. However, this is just an indication of expertise and not support for factors such as outstanding performance or deliberate practice (see above under 1).

All professional interpreters had Swedish as their A-language and English as a C-language according to the language classification of the International Association of Conference interpreters, AIIC.[2] The interpreters with no experience had English as one of their working languages in the translation program. All groups were thus highly proficient in English, but none of them judged themselves to be native or near-native.

3.1.3 *The retrospection procedure*

In accordance with Ivanova's method, the subjects were presented with a transcribed copy of the original speech immediately after the interpreting task. They were asked to go through the speech sentence by sentence and try to recall everything they had thought of while interpreting.

During both the interpreting and retrospection processes, the researcher sat slightly obliquely behind the subject, in order to be close should any technical problems occur, but not too close in order not to disturb the subject during the interpreting process. During retrospection the researcher tried to interfere as little as possible. The retrospection method used here is presented in depth in Englund Dimitrova and Tiselius (2009).

3.1.4 *Transcription and analysis*

The input and output were taped simultaneously on a double-channel recorder and then digitalized as a wav-file. The retrospection interview was taped onto the same tape recorder and then digitalized into a wav-file. The speech, the interpreting and the

2. A-language: the AIIC (International Association of Conference Interpreters) language classification of a language which one masters at mother tongue level and interprets into. B-language: a language other than the interpreter's native language, of which she or he has a perfect command and into which she or he works from one or more of her or his other languages. C-language: languages of which the interpreter has a complete understanding and from which she or he works. http://www.aiic.net/ViewPage.cfm?page_id=199#langclassif (accessed October 14, 2009).

retrospection were all transcribed and coded using Childes software (MacWhinney 1991).

The interpretings were analyzed taking the retrospection as a starting point. Any reported processing problem, instance of monitoring or use of strategy was classified according to Ivanova's categories; although Ivanova only counts a strategy use if it has been coupled to a processing problem or an instance of monitoring, in this study all reported strategies have been counted.

The data were primarily analyzed using *Correspondence Analysis* (CA). CA is an exploratory multivariate statistical method that can be used to visualize the structure of the data (Greenacre 2007: 1). This constitutes a major methodological advantage over looking at percentages calculated from raw frequencies. Consider the difference between the frequencies 100 and 101: in this case, the difference of 1 observed unit is small, since it only amounts to a 1-percent increase. With low frequencies, on the other hand, adding a single observation makes a large difference: if we consider a fairly low frequency such as 5, adding 1 observation would amount to a 20-percent increase, something that makes it nearly impossible to compare percentages from small numbers of observations. CA, on the other hand, takes the total number of observations into account. For technical details, see Greenacre (2007).

The input to a CA analysis is a table, where the rows correspond to observations, the columns correspond to properties of these observations, and the cells contain the frequency of each property per observation (see Table 1). CA is not a hypothesis test (hence, no *p*-values are produced for the CA analyses); instead, the idea is to describe as accurately as possible the structure in the data, primarily through graphical representations called *biplots*.

The biplots can be read as associations between row and column variables, that is, between observations and their properties. For instance, two row variables that are close to each other on the *x*-axis of the biplot are more closely associated than two variables that are further apart. Similarly, distance between variables on the *y*-axis can also be interpreted as associations.

The great advantage of CA is that it allows for simultaneous visualization of the interaction of a number of nominal row and column variables. Thus, the analysis of any one variable will be enlightened by its simultaneous interaction with all the other variables. All the statistical analyses were carried out using the software package R (R Development Core Team 2009); the CA analyses were carried out with the *ca* package in R (Greenacre and Nenadic 2007).

3.2 Results

Below, we present the results of the investigation of processing problems, instances of monitoring and use of strategies among the three different groups of interpreters.

Table 2. Processing problems. Definitions of the headings are given in Appendix 1

Processing problems (PP/)		Long experience interpreters	Short experience interpreters	No experience interpreters
Comprehension (C/)	Perception (P)	8	6	11
	Lexical access in SL (L)	1	0	20
	Syntactic processing (Syn)	1	0	2
	Text integration (TC/integ/)	5	7	16
	Text comprehension (TC/bgkn)	5	3	16
Translation (Tr/)	TL retrieval (TLr)	6	8	24
	Equivalent (eqv)	5	4	4
Simultaneity of tasks (Sim/)	(SL, TL)	7	2	7
	TL delays (Tr.del)	0	0	8
Total		38	30	108

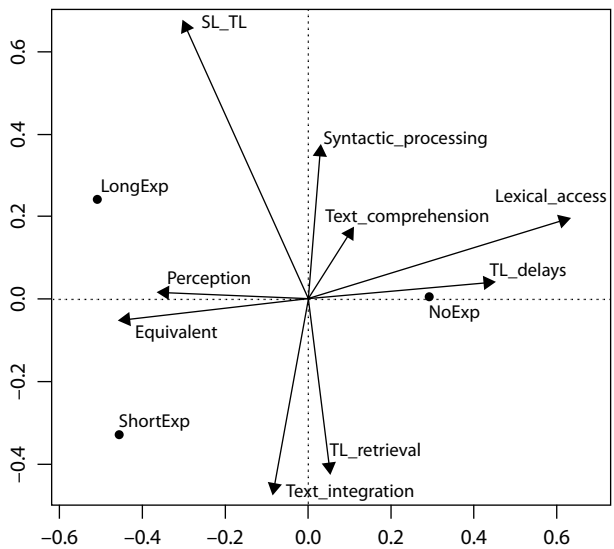


Figure 1. CA biplot of the association between interpreter experience and processing problems. The first axis accounts for 82.8% of the total inertia, the second for 17.2%

3.2.1 *Processing problems*

As can be seen in Figure 1 and Table 2, the largest difference in processing problems is between experienced interpreters and interpreters with no experience. The difference between interpreters with long or short experience is less pronounced. Nevertheless, there are interesting differences in the types of processing problems.

Figure 1 shows a CA plot of the data in Table 2. Most of the variation in table 2 is accounted for by the *x*-axis, and there is a clear difference between long and short

Table 3. Processing problems long experience Ivanova vs. the present study

Processing problem	Ivanova's experts (av. 9 years of interpreting experience)	Long experience interpreters'
P	8	6
L	1	2
Syn	1	2
TC	10	20
Tr	11	16
Sim	7	2
Total	38	48

experience on the one hand, and no experience on the other. All three groups are fairly strongly associated with the *x*-axis judging from their correlations with the *x-axis* (see Table A.4, Appendix 3), which suggests that there is a primary difference between degrees of experience and no experience at all.

Turning to the processing problems, the categories with the highest inertia are *lexical access* (Lexical_access), *target language delays* (TL_delays), *equivalent* (Equivalent), and *source language input rate* (SL_TL) problems. The first two are positively associated with no experience (and negatively associated with short or long experience), equivalents are positively associated with both short and long experience (and negatively associated with no experience), while source/target language problems are positively associated with long experience, but hardly associated with short or no experience at all. The other processing problems have very little explanatory value compared with these four main categories.

However, there is a further distinction between interpreters with short experience and those with long experience. When considering the combined effects of inertia and correlation with the axis, it becomes clear that it is primarily the *source language input rate* (SL_TL) problems that are responsible for the difference between long and short experience. This category is positively associated with long experience on the *y*-axis (and negatively associated with short experience). We find a similar (but weaker) pattern for syntactic processing that appears to help distinguish interpreters with long experience from those with short experience, but it is clearly secondary to the *source language input rate* (SL_TL). Turning to *equivalent*, the *y*-axis reveals that problems with this category appear to be somewhat more frequently associated with short experience than with long experience. See Tables A.4 and A.5 in Appendix 3 for the full numerical output.

In Table 3 our findings are compared with Ivanova's findings. The observed frequencies were compared by means of a Spearman rank sum test, which found no significant difference between the relative ordering of the categories in the two studies ($p > 0.05$). Thus, we can conclude that the figures vary along the same lines, and that, although some variations (for *text comprehension* and *translation*) the figures in the two studies clearly support each other.

Table 4. Processing problems Ivanova's novices vs. short and no experience interpreters in this study

Processing problem	Ivanova's novices (three months interpreting training)	Short experience interpreters	No experience interpreters
P	24	6	11
L	13	0	20
Syn	6	0	2
TC	39	10	32
Tr	21	12	28
Sim	26	2	15
Total	129	30	108

Table 5. Instances of monitoring

Monitoring code (M/)	Long experience interpreters	Short experience interpreters	No experience interpreters
Translation (tr)	16	1	12
Inner speech monitoring (insp)	0	1	1
Time (tm)	1	3	2
Internal commentary (int. com.)	14	9	7
Mood	12	16	39
Id	2	0	5
Total	45	30	24

Table 4 compares Ivanova's novices and the short and no experience interpreters of this study. As above, the categories were compared using Spearman's rank test, and again we found no significant difference between the two studies.

3.2.2 *Instances of monitoring*

In Table 5 it is also clear that the interpreters with long experience differ from the other three groups in types of monitoring. We can assume that the long experience interpreters have enough attention to spare (cf. Gile's effort model 1985) to allow themselves to reflect on the accuracy of the utterance before pronouncing it, while this is not the case for less experienced interpreters.

Figure 2 presents a CA biplot of the data in Table 5. The plot shows a distinction between interpreters with long experience on the one hand, and those with short or no experience on the other. Long experience is associated with a high inertia and is strongly correlated with the x-axis. Short and no experience are more correlated with the y-axis than the x-axis, indicating that they are primarily distinguished from each other on the y-axis. The difference between short and no experience should not be underestimated, since the y-axis accounts for 37.6% of the variance. However, the

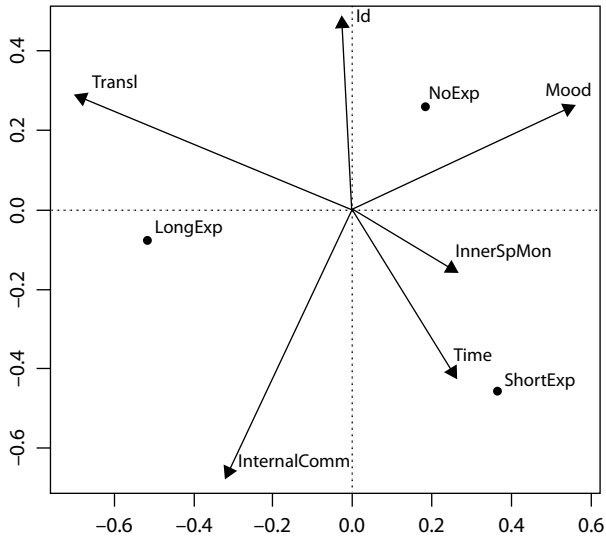


Figure 2. CA biplot of the association between interpreter experience and instances of monitoring. The first axis accounts for 62.4% of the total inertia, the second for 37.6%

primary difference appears to be that between long experience, and short and no experience.

The monitoring observations that have the highest explanatory value (judging from inertia) are *translation* (Transl), *time*, *inner speech monitoring* (InnerSpMon), and *non-analyzed problems* (Id). *Internal commentary* contributes as much as we would expect given its relative frequencies, while *mood* appears to have less explanatory value than its relative frequency would suggest. The strongest contribution to the *x-axis* is made by *translation*, which is positively associated with long experience and negatively associated with the other two categories. There is also some contribution from *mood*, that is, emotive self-evaluations (primarily associated with no experience). On the *y*-axis, short and no experience are distinguished by *internal commentary*, *Id* (non-analyzed problems), and *time*. *Time* and to a certain extent *internal commentary* are positively associated with short experience (and negatively associated with no experience), while *Id* is positively associated with no experience (and negatively associated with short experience).

Thus, *translation* is by far the most frequent monitoring observation for interpreters with long experience, a feature that distinguishes long experience interpreters from the other two groups. The most frequent monitoring observation for interpreters with short experience is *time*, while those with no experience seem to experience *non-analyzed problems* (Id), instances on monitoring that it has not been possible to classify.

See Tables A.6 and A.7 in Appendix 3 for a full numerical output.

Table 6. Instances of monitoring Ivanova's expert vs. the long experience interpreter in the present study

Monitoring code	Ivanova's experts	Long experience interpreters
Tr	18	16
Time	5	0
Insp	8	1
Int. com.	6	14
Mood	9	12
Id	0	2
Total	46	45

Table 7. Instances of monitoring Ivanova's novices vs. the short and no experience interpreters in the present study

Monitoring code	Ivanova's novices	Short experience interpreters	No experience interpreters
Tr	2	1	12
Time	2	1	1
Insp	4	3	2
Int. com.	1	9	7
Mood	7	16	39
Id	0	0	5
Total	16	30	24

The types of monitoring instances of translation also differ, despite the equal numbers for long experience and no experience interpreters. The interpreters with long experience reflect on better ways to interpret a certain utterance, whereas the no experience interpreters try to find a general coherence in the output.

The similarities to Ivanova's findings are not as obvious for interpreters with long experience as they are when comparing processing problems. The two most important instances, viz. translation and mood, are in fact similar; however, whereas Ivanova's experts report both on time and inner speech monitoring, our interpreters with long experience have few and no such instances. Instead, our interpreters with long experience have a number of instances for internal commentary, which Ivanova's interpreters with long experience have considerably less of. However, as with processing, the Spearman rank test found no significant difference between the two studies, which suggests that the observed differences are so small that the two studies can be compared.

Table 7 compares Ivanova's novices to the interpreters with short and no experience in this study. *Mood* is a category that stands out both in Table 7 and Table 6. Simultaneous interpreting seems to be a very emotional experience, both for experienced interpreters and novices, although a little less so for the former. As in the previous cases, we used Spearman's rank test to compare the results of the two studies,

Table 8. Instances of reported strategies

Strategy code	Long experience interpreters	Short experience interpreters	No experience interpreters
Selection (S)	1	1	1
Summarization (SUM)	3	4	0
Restructuring (Rest)	2	3	1
Creative Interpretation (CR)	1	5	9
Overgeneralization (Overgen)	9	5	2
Deletion (D)	18	28	38
Explication (Expl.)	0	0	0
Compromise (ACC)	7	16	12
Total	41	62	63

and the result was not significant, which indicates that also in this case, no overall difference exists between the two studies.

3.2.3 *Strategies*

Table 8 shows the raw frequencies for strategies employed by the groups of interpreters. To assess the correlation of length of experience with individual strategies, we again made use of CA. The result is shown in Figure 3. The *x*-axis accounts for 85.1% of the variance, and the *y*-axis for the remaining 14.9%. On the *x*-axis, there is a continuum from no experience on the right, via short experience near the middle to long experience on the left hand side.

It should be noted that short and long experience groups are closer to each other than either of them is to the no experience group. However, the numerical analysis reveals that long and no experience are primarily responsible for the *x-axis*, while short experience draws up the *y*-axis; thus, it is the opposition between long and no experience that has the greatest explanatory value. Furthermore, the short experience point is close to the origin, indicating that that category has less explanatory value than the other two.

Deletion (D) is the most common strategy and is not strongly associated with any of the groups; hence, we must turn to other strategies to distinguish the interpreter groups.

Interpreters with long experience are strongly associated with *overgeneralization* (Overgen). Interpreters with no experience, on the other hand, are most strongly associated with *creative interpreting* (CR), i.e. guessing. For the interpreters with short experience, *restructuring* (Rest) and *summary* (SUM) are most characteristic, and the latter in particular appears to have some explanatory value with respect to short experience. *Compromises* (ACC) are also somewhat correlated with this group, but have too low inertia to have much explanatory value. Similarly, *selection* (S) plays virtually no role in defining the projection seen in Figure 3.

See Tables A.8 and A.9 in Appendix 3 for a full numerical output.

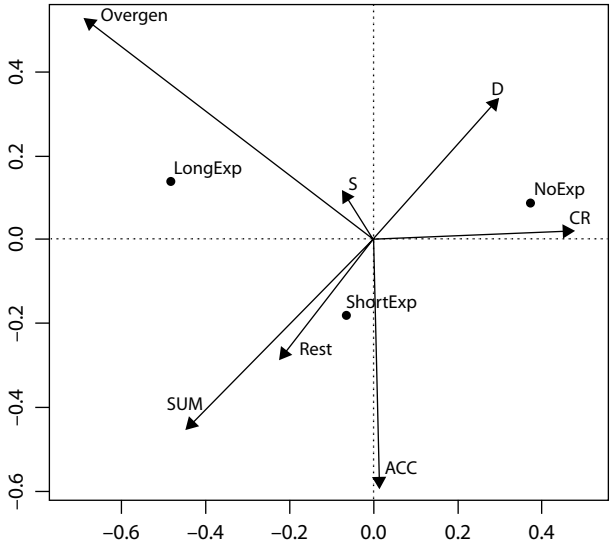


Figure 3. CA biplot of the association between interpreter experience and strategies. The first axis accounts for 85.1% of the total inertia, the second for 14.9%

Strategies have not been compared to the results in Ivanova, since Ivanova only counts strategies in relation to a processing problem, whereas in this study all strategies were counted. Indeed, for the strategies in relation to processing problems, the instances were as low as 13 for interpreters with short experience and 16 for interpreters with long experience. The figures are therefore far too low to make a comparison.

4. Second study: Assessment of the interpreting product

In the second study, we investigated the assessment of the products of the interpreting through ratings. The ratings were done by two groups of raters, professional interpreters (n = 6) and students (n = 6), using Carroll's (1966) scales to rate.

The results of the interpreting assessments were then compared across the groups of interpreters with different levels of experience, with the aim of investigating whether the interpreter's experience made a difference.

4.1 Material and method

The data for the rating study was the same as in the process study: transcriptions of the nine interpretings of the same nine-minute-speech (see Sections 3.1.2 and 3.1.3 above). In an initial, unpublished study the assessment of interpreting from

transcripts and sound-files of the same interpreting were compared and no significant difference was found. Therefore transcripts were used in order to avoid recognition of the interpreters in the study by the interpreter-raters.

4.1.1 *The rating files*

In order to assess the product, all of the transcribed interpreting performances were divided into shorter units, in total 162. These units were then mixed randomly and assembled to six assessment file consisting of 54 units (meaning that each unit got rated twice by different assessors). Each assessment file consisted of randomly mixed units from all the original interpreting performance and without a logical inner order from start to end. These steps were done in order to create an assessment with as even conditions for each assessor as possible. It was assumed that having one rater rate the same interpreter from start to end could possibly distort the rating. The rater could get used to the interpreter's style and this could either improve or impair the scores.

4.1.2 *The raters*

The raters consisted both of experienced interpreters (n = 6) and non-interpreters (n = 6) in order to obtain as broad assessments as possible and to even out differences that might be due to whether the assessor had interpreting experience or not. Gile (1999), for instance, found in one experiment that interpreters tended to be more lenient in their assessment of their peers than non-interpreters. The assessors were trained for the rating immediately prior to the task. Interpreter graders had 8–9 years of interpreting experience and both teaching and testing experience. Non-interpreter graders were university students from various university programs with at least 3 years of university studies.

4.1.3 *The rating session*

Each assessor was given a file of 54 units to grade according to Carroll's scales for intelligibility and informativeness. When assessing intelligibility, the raters only saw the interpreted version; once they had rated the speech for intelligibility, they rated for informativeness by comparing the interpreting with the original speech.

4.1.4 *The scales and the analysis*

The scale for intelligibility is a six-point scale ranging from one to six, where six is the best score. When the interpreting unit is assessed for intelligibility, it is assessed as an independent speech segment of its own. This scale covers assessments of qualities such as fluency, clarity and so forth. In contrast, the scale for informativeness is used to assess the interpreted speech in relation to the source speech, evaluating how much it differs in terms of information from the original speech. This scale covers assessment of fidelity and accuracy. It runs from six to one, with one being the best score. It also has a step 0 to cover explicitation (cf. Englund Dimitrova 2005) where the interpreting contains more information than the original.

These scales were designed to assess the interpreters' rendition from a perspective of general understanding, therefore they can be said to have holistic qualities. The scales are provided in Appendix 3. For a detailed presentation of the methodology of this part of the study, see Tiselius (2009). In the present study, the Friedman rank sum test has been used to test whether a difference exists between the groups of interpreters in terms of their product.

4.2 Results

In this section we present the results for informativeness and intelligibility from the interpreting transcripts.

4.2.1 *Intelligibility and informativeness*

Table 9 gives the scores from the assessment for intelligibility and informativeness of the interpretings. As pointed out in Section 4.1.4 above, a high score for intelligibility means a more comprehensible interpreting product, whereas a low score for informativeness means little difference between source speech and target speech.

To investigate the possible differences between the groups of interpreters, we carried out two Friedman rank sum tests on the ratings for *intelligibility* and *informativeness*, testing the null-hypothesis that there is no difference between the interpreter-experience groups. One test was carried out for *intelligibility* (Friedman $\chi^2_{(df = 2)} = 38.5$, $p < 0.01$), and one for *informativeness* (Friedman $\chi^2_{(df = 2)} = 53.7$, $p < 0.01$). There was a significant difference at the 0.01 level in both cases, indicating that there are differences between the groups somewhere in the data. To pinpoint the between-group differences, we applied a Nemenyi-Damico-Wolfe-Dunn post-hoc test for pairwise comparisons.[3] We found that for *intelligibility*, there is a significant difference between no experience and long experience, and between no experience and short experience. No difference was found between long and short experience with respect to *intelligibility*. For *informativeness*, we found that all three groups are significantly different from each other. Figures 4 and 5 show boxplots of the differences between groups and

Table 9. Mean evaluation scores of intelligibility and informativeness

Interpretings	Intelligibility		Informativeness	
	Non-interpreter grader	Interpreter grader	Non-interpreter grader	Interpreter grader
No experience	3.79	3.16	4.42	5.13
Short experience	5.25	4.88	3.15	3.42
Long experience	5.42	5.11	2.31	2.60

3. R code for this test was provided by T. Galili from http://www.r-statistics.com/2010/02/post-hoc-analysis-for-friedmans-test-r-code/.

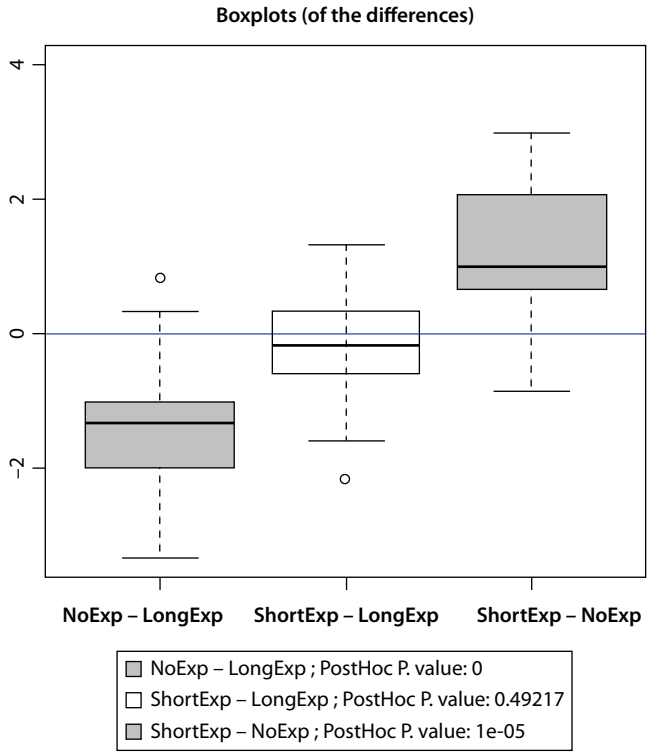


Figure 4. Intelligibility: boxplot of between-group differences. Significant differences are gray. Note the non-significant difference between short and long experience

give the results of the post-hoc tests. The plots show differences between groups: a result close to zero indicates no difference.

The test results suggest that for *intelligibility*, experience vs. lack of experience has an effect, but once a minimum experience threshold is reached, there is no added effect. For *informativeness*, on the other hand, there appears to be a cumulative benefit of experience, so that a little experience is better than no experience, and long experience is better than short experience.

4.2.2 *The effect of raters*

In order to assess the reliability of the results reported above, it was important to quantify the influence that the selection of raters had on the results. As Table 9 above indicates, the scores for non-interpreter (i.e. student) graders are a little more extreme than those produced by experienced graders. Since the Friedman rank sum test employed above is robust and makes very few assumptions about the data, it is a good indicator of any differences between groups. However, it does not take the effect of the raters into consideration.

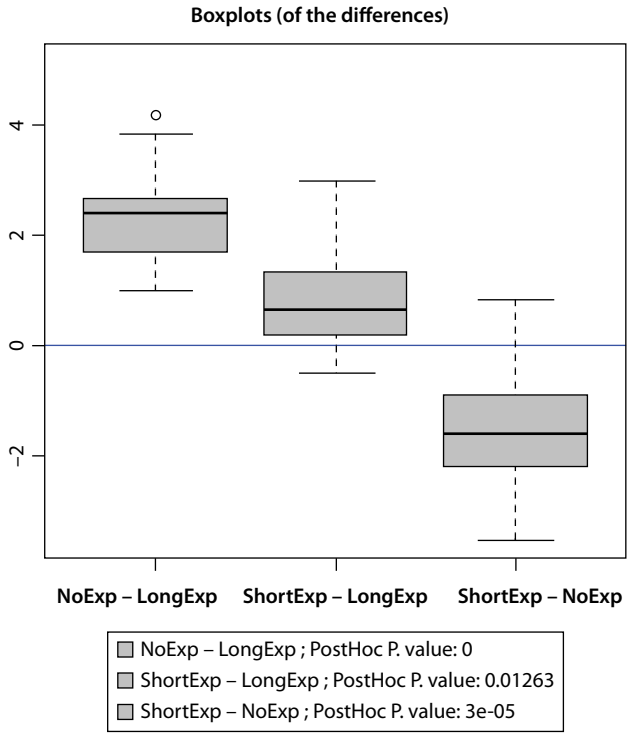


Figure 5. Informativeness: boxplot of between-group differences. Significant differences are gray. All differences are significant, but the largest differences are experience vs. no experience

In order to measure the effect of the raters, a more complex test type known as a mixed effects model was employed. For our purposes, it is sufficient to note that such models allow for complex interactions between multiple factors, while simultaneously incorporating information about group correlations in the data. A full treatment of such models is not possible here; see instead Pinheiro and Bates (2000) or Baayen (2008) for introductions and explanations. The models were fitted with the lme4 package (Bates and Maechler 2009) in R.

The problem to be solved with this approach was as follows: the raters were drawn from two different groups, which might influence the outcome of the rating session. Treating the ratings as coming from a homogeneous group of assessors could potentially lead to misleading results if these rater groups perceived the products in very different ways. For both data sets (viz. *intelligibility* and *informativeness*), the same approach was followed: first, a model was fitted that measured the simple effect of interpreter experience (no, short, long) and rater group (student, interpreter) on the ratings; second, a more complex model measuring the interaction between interpreter experience and rater group was fitted to the same data. All models were found to fit the data well, judging by standard diagnostic procedures.

Overall, the results were congruent with the results presented in the previous section: for intelligibility, we found a statistically significant difference between no experience and long experience ($p < 0.01$), but not between long experience and short experience ($p = 0.23$). For *informativeness*, we found statistically significant differences between no experience and long experience, as well as between long experience and short experience (all $p < 0.01$).[4] This supports the conclusions from the previous sections with respect to the effect of interpreter experience.

In the first, simple model fitted to the data sets, a significant difference between the rater groups was found both for *intelligibility* and *informativeness* (both $p < 0.05$). This was expected, based on the numbers reported in Table 9. The real question, however, was whether this had any systematic effects on the overall outcome of the assessments. The second, more complex model, which measured the interaction between rater group and interpreter experience, found this not to be the case (all $p > 0.05$). Once the interactions were taken into account, the effect of rater group was no longer statistically significant. This result corroborates the impression from Table 9, namely that although the student raters as a group give more extreme assessments than the experienced raters, this has no effect whatsoever on the relative ranking of the products by the three groups of interpreters. In conclusion, there is a difference in magnitude between student and experienced assessors, but the overall trend of the ratings is the same, which suggests that the approach in Section 4.2.1 above was justified.

5. Discussion

The aim of the two studies presented here was to explore the difference (if any) of both process and product between interpreters with no experience, short experience and long experience. This section discusses the result of this study as well as some limitations. Finally, some conclusions regarding the initial research questions are offered.

5.1 Discussion of results

Our study involved relatively few participants; hence, it is difficult to draw exhaustive generalizations from our material. However, until larger studies are conducted, the material allows for certain tentative generalizations. First, the patterns that emerged from our material were relatively clear, and consistent with previous research and theoretical assumptions. Second, the material was gathered under controlled, experimental conditions. This implies that all the participants were tested under the same conditions, which should bring to the forefront any differences due to amount of experience by removing other confounding factors that might otherwise arise.

4. The *p*-values were obtained through MCMC sampling with the *pvals.fnc* function from Baayen (2009).

The experimental set-up has the added benefit of offering compatibility with future studies that employ a similar approach.

With respect to processing problems, we found that the no experience group struggled mostly with *lexical access* and *target language delays*. The experienced interpreters (long and short) were clearly distinct in this matter from those with no experience. For those with short experience, *equivalents* were found to be the most characteristic processing problem. The interpreters with long experience had no difficulties with *delays* or *lexical access*, but experienced some difficulties with *source language input/target language output rate* (SL_TL).

In the case of monitoring a different pattern emerged, where long experience was distinguished from the other two categories; the *amount* of experience appears in other words to play a more active role. For long experience, there was a preference for *translation* (i.e. ascertaining accuracy of translation). Interpreters with short and no experience were grouped together, but could be further distinguished by the fact that short experience was associated with *time* and no experience with *Id*.

In the case of translation strategies, long experience and no experience stood out, whereas interpreters with short experience were fairly close to the overall average. The most common strategy in all groups was *deletion*, which is not surprising considering the fact that interpreters work under a time constraint and a certain amount of deletion comes from compressing the message. Consequently, this deletion could not be used to distinguish the groups. Instead, the most distinctive strategy for interpreters with long experience was *overgeneralization*, which is, using a more general term in the translation (such as "mammal" for "cat"). For interpreters with no experience, *creative interpreting* was the most important strategy, whereas those with short interpreting experience (which was close to the expected average) showed a preference for *summary*.

It was found that the results concerning processing problems in our study supported the findings in Ivanova (1999). For monitoring, the figures were not as unambiguous as for processing problems. With respect to the categories *translation* and *mood*, we could also corroborate Ivanova's results. However, it should be noted that many of the monitoring tendencies may be due to personality or interpreting norms, which may differ according to nationality, language, and interpreting school.

The experience levels of the interpreters participating in the present study and that of Vik-Tuovinen (2006) are not completely similar. However, our results support Vik-Tuovinen's (2006) conclusion that beginners focus on source text and linguistic expression, since the subjects with no experience in the present study struggle with problems of *lexical access*. Among the interpreters with extensive experience in the present study, the monitoring category *translation* dominated, and this might very well mirror Vik-Tuovinen's experienced interpreters, who focused more on situational factors than less experienced interpreters.

Liu (2001) found that interpreters with extensive experience monitored the output better. Although her test subjects were not fully comparable in terms of experience to the subjects participating in the present study, Liu's results were supported in

our study of the interpreting process, interpreters with long experience make more use of monitoring strategies, and in particular *translation* and *internal comments*.

When it comes to the analysis of the product, the data for *intelligibility* and *informativeness* were tested using the Friedman rank sum test and the Nemenyi-Damico-Wolfe-Dunn post-hoc test. The results clearly indicate that experience results in higher *intelligibility* once a minimum level of experience has been attained, but that there is no added benefit of longer experience beyond this minimum. For *informativeness*, on the other hand, longer experience has a cumulative effect, so that a short experience leads to better scores than no experience, and long experience gives better scores than short experience. Through mixed effects models, it was demonstrated that rater experience did not influence the overall trend of the assessment of the product.

The fact that *Intelligibility* is not a distinguishing factor between interpreters with short and long experience is not surprising considering the fact that interpreters with training and some professional experience have been taught how to produce a comprehensible message. To find differences in the interpreting product between interpreters with long and short experience, it is necessary to look beyond *intelligibility*, to *informativeness*.

5.2 Conclusion

We have demonstrated that the impact of experience in interpreting is dependent on the task or problem at hand. Both interpreting experience and general background knowledge are important for the solution of processing problems in different situations. In all cases, it would seem that some minimum of experience is desirable, but only in some areas did we find a difference between the long and short experience groups.

It seems there is a clear difference between the dichotomy experience vs. lack of experience both in process and product. It should be noted that a major difference in terms of experience between the student group (no experience) and the two groups of interpreters with some degree of experience, is that the two groups of interpreters both had experience interpreting at the European Parliament and the speech chosen was from the European Parliament. This might give them a double edge not only in terms of simultaneous interpreting technique, but also in knowledge of the European institutions. Thus, the groups of experienced interpreters had experience in that field and perhaps a certain amount of expertise as well.

Degrees of experience (short and long) proved a less profound opposition in our material, although some differences could be seen in the preferences of strategies and perceived problems of the long and short experience groups. The exceptions to this were monitoring and *informativeness*, where long experience appears to have a larger impact (and give more of an advantage) than in the other cases. It should also be stressed that we still do not know for sure whether the interpreters with long experience are experts or not. This is a limited study and all criteria for studying expertise

(see Section 1) have not been covered here. Expertise must necessarily be studied through proxy indicators. Although the indicators studied could be expanded, it can be assumed that these very experienced interpreters have at least some degree of expertise. When we combine the results of the two studies we see that for strategies, long experience interpreters report more instances of *overgeneralization*, whereas short and no experience interpreters report more *deletion* and *creative interpreting*. Since subjects rate products from interpreters with long experience higher, this can indicate that the interpreting is perceived to be closer to the original when overgeneralization is employed, compared with deletion or creative interpreting.

To sum up, certain processing problems, instances of monitoring and strategies may indeed tell us something about expertise. They may indicate which components of the performance should be studied in detail to find the superiority Ericsson (1991) encourages us to look for. It may very well be that monitoring and *informativeness* are the components that make the expert performance superior. The challenge ahead is to dig deeper into these differences in order to understand and define expertise in interpreting.

References

Anderson, L. 1979. "Simultaneous interpretation: Contextual and translation aspects." (Master's thesis. Department of Psychology, Concordia University, Montreal, Canada.)

Baayen, R. H. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Baayen, R. H. 2009. *languageR: Data Sets and Functions with "Analyzing Linguistic Data: A Practical Introduction to Statistics." R Package Version 0.955*. http://CRAN.R-project.org/package=languageR. (Accessed April 19, 2010.)

Bates, D. and Maechler, M. 2009. *lme4: Linear Mixed-Effects Models using S4 classes. R Package Version 0.999375-32*. http://CRAN.R-project.org/package=lme4. (Accessed April 19, 2010.)

Bühler, H. 1986. "Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters." *Multilingua* 5 (4): 231–235.

Dillinger, M. 1989. "Component processes of simultaneous interpreting." (Doctoral dissertation, McGill University, Montreal, Canada.)

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process*. Amsterdam/Philadelphia: John Benjamins.

Englund Dimitrova, B. and Tiselius, E. 2009. "Exploring retrospection as a research method for studying the translation process and the interpreting process." In *Methodology, Technology and Innovation in Translation Process Research* [Copenhagen Studies in Language 38], F. Alves, S. Göpferich and I. Mees (eds), 109–134. Copenhagen: Samfundslitteratur.

Ericsson, K. A. (ed.). 1996. *The Road to Excellence: The Acquisition of Expert Performance in Arts and Science, Sports and Games*. Mahawa, N.J.: Lawrence Erlbaum.

———. 2000. "Expertise in interpreting: An expert-performance perspective." *Interpreting* 5 (2): 189–222.

———. 2008. "An introduction to the *Cambridge Handbook of Expertise and Expert Performance*." In *The Cambridge Handbook of Expertise and Expert Performance*, K. A. Ericsson, N. Charness, P. J. Feltovich and R. R. Hoffman (eds), 3–20. New York: Cambridge University Press.

Ericsson, K. A. and Smith, J. 1991. "Prospects and limits of the empirical study of expertise: an introduction." In *Toward a General Theory of Expertise: Prospects and Limits*, K. A. Ericsson and J. Smith (eds), 1–38. Cambridge. Cambridge University Press.

Gile, D. 1985. "Le modèle d'efforts et l'équilibre en interpretation simultanée." *Meta* 32 (2): 164–169.

———. 1999. "Variability in the perception of fidelity in simultaneous interpretation." *Hermes – Journal of Linguistics* 22: 51–79.

———. 2003. "Quality assessment in conference interpreting: Methodological issues." In *La evaluación de la Calidad en interpretación: Investigación*, A. Collados Aìs, M. Fernández, A. Sánchez and D. Gile (eds), 109–123. Granada: Editorial Comares.

Gerver, D. 1971. "Aspects of simultaneous interpretation and human information processing." (Doctoral dissertation, Oxford University, UK.)

Greenacre, M. 2007. *Correspondence Analysis in Practice*. 2nd edition. Boca Raton: Chapman and Hall/CRC.

Greenacre, M. and Nenadic, O. 2007. *ca: Simple, Multiple and Joint Correspondence Analysis. R package version 0.21.*

Ivanova, A. 1999. "Discourse processing during simultaneous interpreting: An expertise approach." (Doctoral dissertation, University of Cambridge, UK.)

———. 2000. "The use of retrospection in research on simultaneous interpretation." In *Tapping and Mapping the Processes of Translation and Interpretation*, S. Tirkkonen-Condit and R. Jääskeläinen (eds), 27–52. Amsterdam/Philadelphia: John Benjamins.

Jääskeläinen, R. 2010. "Are all professionals experts? Definitions of expertise and reinterpretation of research evidence in process studies." In *Translation and Cognition,* G. M. Shreve and E. Angelone (eds), 213–227. Amsterdam/Philadelphia: John Benjamins.

Kurz, I. 2002 and 1993. "Conference interpretation: Expectations of different user groups." *The Interpreters' Newsletter* 5: 13–21. Reprinted in: *The Interpreting Studies Reader,* F. Pöchhacker and M. Shelsinger (eds), 312–324. London/New York: Routledge.

Liu, M. 2001. "Expertise in simultaneous interpreting: A working memory analysis." (Doctoral dissertation, University of Texas at Austin, Texas.)

———. 2008. "How do experts interpret? Implications from research in Interpreting Studies and cognitive science." In *Efforts and Models in Interpreting and Translation Research. A Tribute to Daniel Gile*, G. Hansen, A. Chesterman and H. Gerzymisch-Arbogast (eds), 159–177. Amsterdam/Philadelphia: John Benjamins.

MacWhinney, B. 1991.*The CHILDES Project: Tools for Analyzing Talk*. Hillsdale, N.J.: Lawrence Erlbaum.

Moser, B. 1978. "Simultaneous interpretation: A hypothetical model and its practical application." In *Language Interpretation and Communication*, D. Gerver and H. W. Sinaiko (eds), 353–368. New York: Plenum Press.

Moser-Mercer, B. 1996. "Quality in interpreting: Some methodological issues." *Interpreters' Newsletter* 7: 43–55.

———. 2000. "The rocky road to expertise in interpreting: Eliciting knowledge from learners." In *Translationswissenshaft*, M. Kadric, K. Kaindl and F. Pöchhacker (eds), 339–352. Tübingen: Stauffenburg Verlag.

Moser-Mercer, B., Fraunfelder, U. H., Casado, B. and Kunzli, A. 2000. "Searching to define expertise in interpreting." In *Language Processing and Simultaneous Interpreting. Interdisciplinary Perspective*, B. Englund-Dimitrova and K. Hyltenstam (eds), 1–21. Amsterdam/Philadelphia: John Benjamins.

Ng, B. C. 1992. "End users' subjective reaction to the performance of student interpreters." *The Interpreters' Newsletter Special Issue* 1: 35–41.

Pinheiro, J. C. and Bates, D. 2000. *Mixed-Effects Models in S and S-plus*. New York: Springer.

R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. (Accessed April 19, 2010.)

SCIC. 2007. *Customer Satisfaction Survey.*

SCIC. 2010. *Customer Satisfaction Survey*.

Shlesinger, M. 1997. "Quality in simultaneous interpreting." In *Conference Interpreting: Current Trends in Research*, Y. Gambier, D. Gile and C. Taylor (eds), 123–131. Amsterdam/Philadelphia: John Benjamins.

Seleskovitch, D. 1965. *Colloque sur l'enseignement de l'interprétation.* Geneva: AIIC (Association Internationale des Interprètes de Conference).

Tiselius, E. 2009. "Revisiting Carroll's Scales." In *Testing and Assessment in Translation and Interpreting Studies* [ATA Monograph Series], C. Angelelli and H. Jacobson (eds), 95–121. Amsterdam/Philadelphia: John Benjamins.

Vik-Tuovinen, G.-V. 2006. *Tolkning på olika nivåer av professionalitet*. Vasa: Acta Wasaensia.

Vuorikoski, A. R. 1998. "User responses to simultaneous interpreting." In *Unity in Diversity? Current Trends in Translation Studies,* L. Bowker, M. Cronin, D. Kenny and J. Pearson (eds), 184–97. Manchester: St. Jerome.

Zwischenberger, C., Pöchhacker, F. and Kurz, I. 2008. "Quality and role: The professionals' view." *Communicate! A Webzine of The International Association of Conference Interpreters.* http://www.aiic.net/ViewPage.cfm/page3044.htm. (Posted September 5, 2008; updated September 8, 2008; accessed May 3, 2010.)

Appendix 1

Table A.1 Processing problems. Ivanova (1999)

Processing problems (PP/)		
Comprehension (C/)	Perception (P)	Problems with hearing
	Lexical access in SL (L)	Failure to access meaning of a SL chunk, which has been identified as familiar
	Syntactic processing (Syn)	Failure to recognize syntax patterns
	Text integration (TC/integ/)	Difficulties in constructing a coherent representation for SL chunks
	Text comprehension (TC/bgkn)	Comprehension difficulties due to lack of background knowledge
Translation (Tr/)	TL retrieval (TLr)	Problems in rendering a SL chunk in TL
	Equivalent (eqv)	Problems in selecting an appropriate equivalent when there is a choice
Simultaneity of tasks (Sim/)	(SL, TL)	Problems due to high SL input rate in relation to interpreter's own output rate
	TL delays (Tr.del)	Delays in TL product due to translation

Table A.2 Monitoring observations. Ivanova (1999)

Monitoring (M/)	
Translation (tr)	Ascertaining accuracy of translation at the conceptual level against an ST representation
Inner speech monitoring (insp)	Verification of the TL message against TL rules prior to articulation
Time (tm)	Awareness of the timing relative to the TL production
Internal commentary (int. comm.)	Affective commentary to ST/ST producer
Mood	Emotive self-evaluation of performance
Id	Non-analyzed problems

Table A.3 Strategies. Ivanova (1999)

Strategy code (SC/)	**Brief description**
Selection (S)	Selection of one SL chunk for further processing because it is more informationally or pragmatically salient
Summarization (SUM)	Rendering the gist of a SL segment
Restructuring (Rest)	Changing the original syntactic structure of a SL segment (usually by transposing clauses or segments within the clause), in order to improve the expression in TL, or anticipation of problems
Creative Interpretation (CR)	Compensating for missing information by guessing on the basis of previous knowledge
Overgeneralization (Overgen)	Rendering an aspect of ST (e.g. a lexeme) by selecting a more abstract, hence less committing representation in TL
Deletion (D)	Omitting SL chunks because of time constraints without reference to their semantic or pragmatic role
Explication (Expl.)	Explicitly expressing information implied in the ST
Compromise (ACC)	Lower the acceptability standards for a TL production in order to minimize processing costs

Appendix 2

Original speech

Commissioner Byrne on the Tobacco Products Directive, held in the European Parliament May 14, 2001. The text has been slightly edited.

Mr. President, honorable Members,
It is with great pleasure that I reply to this debate on the Committee report concerning the Tobacco Products Directive. As you are all well aware, this directive is the culmination of several years of hard work by the Commission, and has been developed on the basis of the best scientific advice.

We have developed and proposed this directive, which we consider a qualitative step forward in protecting public health. This work has been done within the legal framework for completion of the internal market. The directive before you today will represent a significant improvement on our current legislative position and fill many of the gaps, which have made the current rules ineffective.

I need hardly remind Parliament of the level of public damage to the health of our population from tobacco. This damage exceeds many other threats to public health that are far more widely debated. The World Health Organization, for instance, estimates tobacco-related deaths in the Community to exceed half a million a year, and this number does not even include the number of smokers who suffer from serious and debilitating illnesses, such as emphysema, heart failure, TB, lung and liver cancer. Because of this major impact on human health, tobacco control emerges as the most import tool in reducing death and disease in our midst.

I must here compliment Parliament and in particular the rapporteur, Mr. Maaten. As you all know, the amendments he has tabled have substantially improved and completed the directive. The Commission was happy to accept many of these amendments. Mrs. Maes went as far in her contribution to Parliament as to make a statement that this was a textbook example of cooperation between the three great legislative institutions of the European Union. I would echo that and go even further and say that this issue is worthy of close examination by any serious student of the European Union or its legislative procedure. Both the advertising directive that I will be presenting to you soon, and this particular tobacco directive on which you will vote today, are worthy of such an examination. The directive provides a clear example of how all the institutions can work well together, both in formulating policy, and in bringing that policy forward in the form of valuable legislation. And all this has been done within the confines of the European legislation from the Treaty of Rome and onwards. Also, it will surely calm anybody who fears that the European institutions are becoming too powerful.

This directive recasts several existing directives from the late 1980s. We have used this proposal to update and complete existing instruments on the basis of scientific developments. First of all I should like to mention the setting of a reduced tar ceiling in cigarettes and the creation for the first time of a ceiling for nicotine and carbon monoxide. These ceilings will directly attack the agents in cigarettes responsible for cancer, addiction, and cardiovascular disease. I wish to point out that we are not, however, aiming at producing a safe cigarette, since, as everybody knows, such a product cannot exist.

Importantly, the directive now applies to all imported products and to all products manufactured in the Community. We will no longer export any products that we consider

inappropriate for our citizens. This will also eliminate the risk of unsafe exports coming back onto the Community market through smuggling.

The tobacco industry has tried to use the spectre of job losses to threaten this measure, and these cynical efforts have made me particularly upset. The same threats were also expressed when the original tobacco directives were adopted in the late 1980s. These threats were not true then and they are not true now. There is no room for double standards for our own citizens or others, when we are talking about addictive and deadly products.

Another innovation in the text of this directive is to require that the tobacco producers declare what additives they put in tobacco products – potentially the single most important measure contained in this legislation. Currently we do not know what these additives are in the Member States or, more importantly, why they are added to the product. Manufacturers and importers in future will be obliged to declare these constituents.

As a result of an amendment from this Parliament on the 24th of July, 1999, the introduction of colour photos on product warnings is to be introduced for the first time. This will greatly increase their impact. In recent years, research has shown the importance of avoiding misleading product descriptors on tobacco products, because such descriptors may mislead the consumer into believing that one product is safer than another. This directive introduces a requirement not to use such misleading terms on tobacco products. As such it aims to protect smokers and non-smokers alike from misleading and dangerous descriptors such as light, low tar, double filtered, ultras, etc.

In a previous declaration to the Parliament, I undertook to consult tobacco experts. Richard Peto, who is statistician and epidemiologist at Magdalena College of the University of Loughborough reported that BAT, British American Tobacco, recently gave more than £3 million to sponsor Nottingham University. Now, every cigarette sold makes a profit for the company of about 3p. To get its donation money back, which we have to assume BAT wants to do, it has to sell 100 million extra cigarettes. Every million cigarettes causes about one death. So to break even, they've got to sell enough cigarettes to cause about 100 deaths. The money, incidentally, was given to fund a professorship in corporate responsibility.

I can also inform you today that in the coming weeks, I hope to meet my promise to present a new proposal on tobacco advertising and sponsorship. This will replace the directive annulled by the Court of Justice on the 18th of October last year.

I also regard the Community's legislative activity as a complement to the legislation work under way in the World Health Organization Framework Convention on Tobacco Control. I welcome the presence of an observer from the European Parliament in the Community delegation, which early this month met in Geneva during the second round of negotiations for this convention.

Finally, I should like to pay tribute in particular to Mr. John Ryan and the significant contribution he has made to this issue over the years.

I am convinced that it will be seen in coming years that this directive is a keystone of our efforts to reduce the damage to health caused by smoking.

Mr. President, honorable Members, thank you for your attention. (1093)

Appendix 3

Table A.4 CA output for the biplot in Figure 1 (columns), giving relative frequencies (‰), inertia, correlation with the *x*- and *y*-axes, and contribution to the *x*- and *y*-axes per factor

ProcProblem	RelFreq	Inertia	X-cor	X-contrib	Y-cor	Y-contrib
Prc	142	101	999	122	1	0
L__	119	324	981	384	19	36
Sy_	17	23	33	1	967	127
Txt_n	159	41	133	7	867	208
Txt_c	136	13	661	10	339	25
TL_r	216	31	73	3	927	165
Eqv	74	162	997	195	3	3
SL_	91	147	492	87	508	434
TL_d	45	159	998	192	2	2

Table A.5 CA output for the biplot in Figure 1 (rows), giving relative frequencies (‰), inertia, correlation with the *x*- and *y*-axes, and contribution to the *x*- and *y*-axes per factor

Experience	RelFreq	Inertia	X-cor	X-contrib	Y-cor	Y-contrib
LnE	216	381	816	375	184	409
ShE	170	300	662	239	338	590
NEx	614	320	1000	386	0	1

Table A.6 CA output for the biplot in Figure 2 (columns), giving relative frequencies (‰), inertia, correlation with the *x*- and *y*-axes, and contribution to the *x*-and *y*-axes per factor

Monitoring	RelFreq	Inertia	X-cor	X-contrib	Y-cor	Y-contrib
Trn	206	324	906	471	94	81
ISM	14	49	828	65	172	22
Tim	43	103	390	64	610	167
InC	213	228	267	98	733	445
Mod	475	213	884	302	116	65
Id	50	83	5	1	995	220

Table A.7 CA output for the biplot in Figure 2 (rows), giving relative frequencies (‰), inertia, correlation with the *x*- and *y*-axes, and contribution to the *x*- and *y*-axes per factor

Experience	RelFreq	Inertia	X-cor	X-contrib	Y-cor	Y-contrib
LnE	319	418	977	656	23	25
ShE	213	351	392	220	608	567
NEx	468	231	335	124	665	408

Table A.8 CA output for the biplot in Figure 3 (columns), giving relative frequencies (‰), inertia, correlation with the *x*- and *y*-axes, and contribution to the *x*- and *y*-axes per factor

Strategy	RelFreq	Inertia	X-cor	X-contrib	Y-cor	Y-contrib
S	18	5	707	5	293	11
SUM	42	198	847	197	153	203
Rst	36	53	775	48	225	79
CR	90	181	1000	213	0	0
Ovr	96	427	906	454	94	267
D	506	87	819	84	181	105
ACC	211	50	4	0	996	333

Table A.9 CA output for the biplot in Figure 3 (rows), giving relative frequencies (‰), inertia, correlation with the *x*- and *y*-axes, and contribution to the *x*- and *y*-axes per factor

Experience	RelFreq	Inertia	X-cor	X-contrib	Y-cor	Y-contrib
LnE	247	471	923	511	77	242
ShE	373	103	108	13	892	613
NEx	380	427	950	476	50	144

Appendix 4

Table A.10 Scale of intelligibility as adapted for the present study

6.	The interpreting is perfectly clear and intelligible. Like ordinary spoken Swedish with few if any stylistic infelicities.
5.	Generally clear and intelligible but with minor grammatical or stylistic peculiarities or unusual word choices, though nothing that hampers the understanding.
4.	The general idea is intelligible, but full comprehension is interfered with by poor word choice, poor style, unusual words and incorrect grammar. The Addressee will have to make an effort to understand the utterance.
3.	Masquerades as an intelligible utterance, but is actually more unintelligible than intelligible. Nevertheless, the idea can still be comprehended. Word choices, syntactic arrangements, and expressions are generally unusual, and words crucial to understanding have been left out.
2.	Almost completely unintelligible. Although it does not seem completely nonsensical and the Addressee may, with great effort, discern some meaning.
1.	Totally unintelligible and completely without meaning.

Table A.11 Scale of informativeness as adapted for the present study

6.	Reading the original changes the whole understood meaning. (6 should be given when reading the original completely changes the meaning that the interpreting gave).
5.	Reading the original clarifies the understood meaning. The original's differences in syntax, words and phrases alter the listener's impression of the meaning to some extent.
4.	Reading the original gives some additional information about syntax and words. It can also clarify minor misunderstandings in the interpreting.
3.	By correcting one or two meanings, mainly on word level, the reading of the original gives only a minor difference in meaning compared to the interpreting.
2.	No new meaning is added through reading the original, neither at the word level nor at the grammatical level, but the Addressee is somewhat more confident that s/he really comprehends the meaning intended.
1.	No new meaning is added by the original, nor is the Addressee's understanding of the interpreting increased.
0.	The original contains less information than the interpreting.

Developing professional thinking and acting within the field of interpreting

Gun-Viol Vik-Tuovinen
University of Vaasa, Finland

The aim of this article is to contribute to the discussion about how professional thinking and acting is developed and spread within the field of interpreting. As a theoretical starting point for the discussion, the central factors affecting strategies that interpreters can and do use in performing interpreting tasks are introduced. As a second basis for the discussion the concept of practical and tacit knowledge is discussed. Three empirical studies are then presented. The first one deals with how students of interpreting and interpreters select strategies during the performance of an interpreting task. The other two concentrate on practical and tacit knowledge within the field of interpreting.

Ideally, interpreting performance is based on each situation, makes good use of the individual interpreter's knowledge and mental capacity and respects prevailing norms. The empirical studies presented show the development of thinking and acting related to interpreting among interpreting students and interpreters. The studies indicate that students are aware of and have integrated essential issues related to professional thinking and acting. However, among students and practising interpreters, as well as within the whole field of interpreting, there still is a need to both develop and to transfer professional thinking and acting.

Keywords: interpreter training, strategy, norm, practical knowledge, tacit knowledge

1. Introduction

Today, basic interpreting proficiency can commonly be achieved through training. Most interpreters entering the field of conference interpreting have some kind of prior training, and an increasing number of community interpreters have received training as well. Initially, interpreting students acquire elementary knowledge of interpreting. This consists of the knowledge required to perform routine tasks in

a relatively predictable and stable setting. Gradually, the training comes to include more demanding interpreting tasks, and ideally the students proceed from elementary knowledge to advanced knowledge. Advanced knowledge is based on rules that must be adapted to a particular situation and must be learned in real situations or through authentic examples testing the knowledge held (cf. Rolf 1991: 104–106). Interpreters' thinking and acting are primarily based on knowledge of the situation and circumstances involved in the assignment, their own knowledge and mental potential, understanding of norms, and the knowledge they have gained by contact with the performance of interpreting tasks.

The focus of this article is on the knowledge and skills needed for interpreting. The article is based on the author's research on interpreting at different levels of proficiency (Vik-Tuovinen 2006) and on the concepts of practical and tacit knowledge. Inspiration for the studies on practical and tacit knowledge was drawn from a Swedish project on the tacit knowledge of interpreters in 2008 and 2009 (cf. Holmgren 2009).

In Section 2, an overview of the factors laying a foundation for the interpreter's performance and problem solving is presented. In Section 3, the concepts of practical and tacit knowledge are discussed. Three empirical studies are presented in Sections 4–6. All the studies were conducted in Finland. In the first study, differences and similarities between students and practising interpreters are approached using the retrospection method to elicit information about the interpreting process. The focus of this study is on which phases of the interpreting process new and advanced students and practising interpreters emphasise in their retrospective comments on an interpreting task. The development of professional thinking and acting is further discussed on the basis of results of a questionnaire study with interpreting students and interpreters as respondents. The third study is an interview study with practising interpreters, which aims to follow up on and deepen the understanding of the results of the questionnaire study.

2. Interpreting strategies as part of the interpreter's thinking and acting

Interpreters need to apply relevant strategies to perform their assignments successfully, and the use of proper strategies is an essential element of interpreters' thinking and acting. When applying strategies, interpreters utilise their personal knowledge, their knowledge related to the current interpreting assignment and their understanding of good interpreting in general. The interpreting strategies that interpreters can and will apply depend on several factors. Three main categories of factors are *situational factors*, *the interpreter's knowledge and mental potential* and *norms*. These categories of factors are not autonomous and may also be interdependent.

Interpreting strategies can be seen as procedures activated by interpreters in order to manage an interpreting assignment and to solve or head off problems (cf. Lörscher 1991: 76; Kalina 1998: 113–125; Vik-Tuovinen 2006: 39). Strategies may be applied

before, during and after the assignment. They can be related to the assignment as whole, for instance strategies related to preparations for the coming assignment, or to a single linguistic problem, as when selecting a correct expression to use as an equivalent for an expression used by the speaker. When a strategy is triggered by a problem or a potential problem, application of the strategy is conscious, otherwise it may be automatic and unconscious (cf. Jääskeläinen and Tirkkonen-Condit 1991: 90–91, 94; Vik-Tuovinen 2006: 144–146). Conscious application of a strategy can be explained and verbalised. In Translation Studies, thinking-aloud methods have been applied to gain information about the translation process (cf. Krings 1986; Jääskeläinen and Tirkkonen-Condit 1991; Lörscher 1991; Englund Dimitrova 2005), and in Interpreting Studies, retrospection has been used as a corresponding method for eliciting information (cf. Ivanova 1999; Mead 2002; Bartłomiejczyk 2004; Vik-Tuovinen 2006).

The main categories of factors affecting interpreters' strategies – situational factors, the individual interpreter's knowledge and mental potential and norms – can further be divided into subcategories. In Figure 1, the main categories and central subcategories are presented in an overview. The overview was originally created by the author to give a structure to the factors affecting strategies that interpreters apply when performing an interpreting task (Vik-Tuovinen 2006: 33–124) and as a frame for essential issues that should be included in interpreter training. The overview is based on theory and research within the fields of Translation and Interpreting Studies.

Situational factors mentioned in Figure 1 consist of the participants of the event being interpreted, the purpose of the communication, source text characteristics and the mode of presentation of the source text, and the event setting (cf. Gerver 1969/2002; Anderson 1978: 219–228; Namy 1978: 29; Poyatos 1987; Shlesinger 1989: 14–69; Pöchhacker 1994: 65–75; Alexieva 1997/2002: 220–230). The interpreter will use strategies on the basis of an understanding of the situational factors and the ability to adapt to these factors.

An individual interpreter's knowledge and mental capacity can be considered as laying the foundation for the performance of an interpreting task. Essential parts of the interpreter's knowledge are general knowledge and knowledge of the subject field and terminology for the current assignment (Vik-Tuovinen 2003; Sunnari 2004: 57–58). Sufficient language skills are a prerequisite for interpreting. Knowledge of interpreting, including techniques for performing simultaneous and/or consecutive interpreting, is of essential importance for successful performance of any interpreting assignment.

The norms affecting interpreting strategies can be divided into ethical norms, inter-textual norms related to the relationship between source text and interpretation, intra-textual norms related to interpretation as received by a person who does not understand the source language, and instrumental norms related to interpreting as a service for clients (Wadensjö 1992; Toury 1995: 55–56; Chesterman 1997; Shlesinger 1997; Garzone 2002; Pöchhacker and Shlesinger 2002: 295–298; Vik-Tuovinen 2006: 82–119).

The more successfully interpreters master all the categories presented in Figure 1, the more prerequisites for a successful performance they have in hand. The interpreter

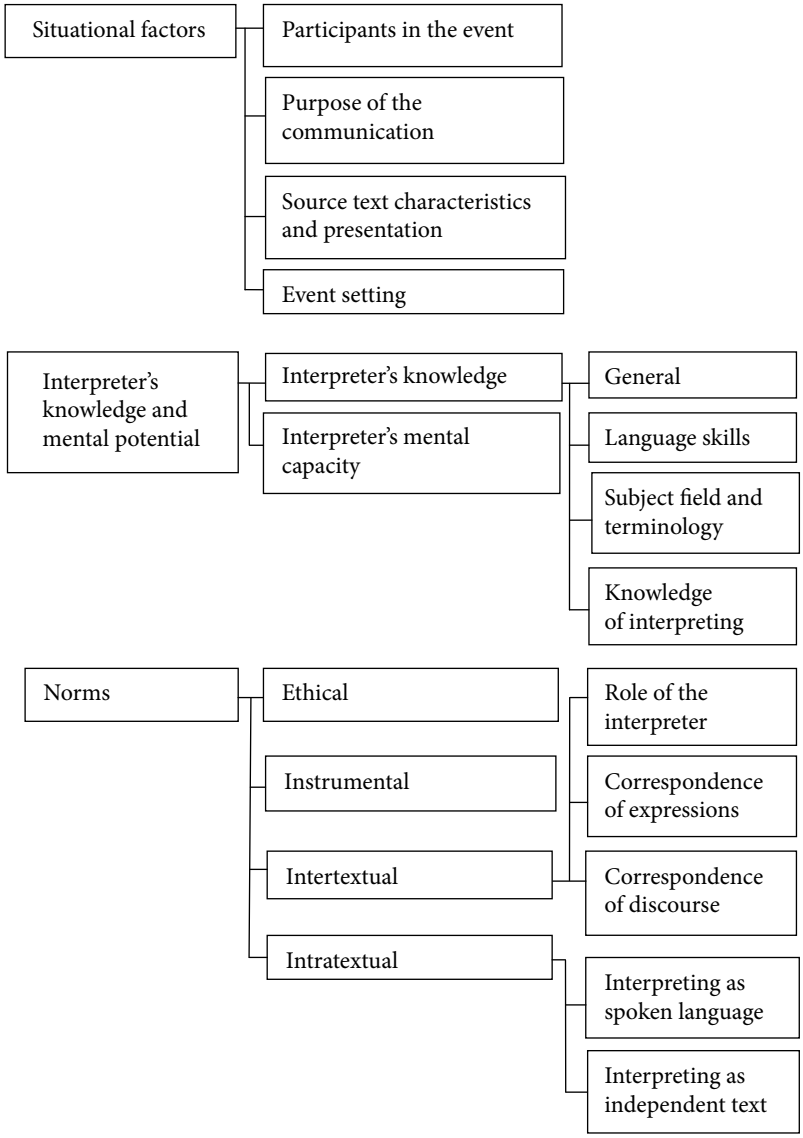


Figure 1. Factors affecting interpreting strategies

may only influence the situational factors to a certain degree, mostly having no choice but to adapt to them. When it comes to the interpreters' own knowledge and mental potential, they themselves can improve them by learning more and developing their mental capacity. An interpreter's knowledge and understanding of norms can also be improved. Training plays an essential role, but learning will continue throughout the whole career of the practising interpreter.

3. Practical and tacit knowledge as part of the interpreter's thinking and acting

Professionals from many fields generally have access to knowledge acquired through training and written sources. But they also have practical knowledge related to the profession. This knowledge has developed through practice and personal experience and through interaction with others involved in the field. It includes that which a person has internalised in terms of competence, skills and values (Rolf 1991: 27–28). This kind of knowledge is also called tacit knowledge (Polanyi 1966).

The concepts of practical knowledge and tacit knowledge are often understood to overlap or be synonymous (cf. Tuomola and Airila 2007). Professor Svenaeus, a Swedish expert on the subject, affirms that the terms *practical knowledge* and *tacit knowledge* can be used synonymously, but he himself prefers to call the knowledge practical, as the knowledge is not in fact tacit, but can be verbalised and reflected upon (Holmgren 2009: 11–12).

According to Gustavsson, practical knowledge has been described as knowledge related to skills, to *know-how*, and it is related to a certain context or situation. Additionally, he mentions that practical knowledge has to do with reflection on one's own action. (Gustavsson 2000: 108–109, 128.) In Johannesson and Rolf (1989: 31), tacit knowledge is considered to be knowledge acquired through a person's direct experience of whatever his or her knowledge covers, that is acquired in the same environment where it will be used.

Practical and tacit knowledge can thus be gained by personal experience and reflection. According to Polanyi and Prosch (1975: 32) this knowledge can be transmitted to other people in the same way knowledge is transmitted by adults to children, or the way medical students are taught scientific skills and acquire expert knowledge at the bedside and in the laboratory.

Practical and tacit knowledge related to interpreting should thus also be transferable in real situations or during training adapted to the needs of authenticity. The context and situations must also provide opportunities for personal experience and reflection. Authentic interpreting exercises, observations of real interpreting situations and interpreting practice are, according to a survey among teachers of interpreting at the universities in Finland, regularly integrated into the training of future interpreters (Vik-Tuovinen 2008). Students of interpreting are thus introduced to situations where practical and tacit knowledge can be transferred.

As practical knowledge can be reflected on and put into words, it is possible to discuss it with others and transfer it to others. It can be transferred in teaching and when people relate or discuss their own experiences, and it can be written down and thus made available to the general public.

Rolf (1991: 27) states that tacit knowledge can be categorised according to function, source and contents. Table 1 displays how verbalised practical and tacit knowledge has been, or can be, made available within the field of interpreting. The categorisation made by Rolf has been slightly simplified and adapted to better describe the practical

Table 1. The function, source and content of practical and tacit knowledge

Function	Source	Contents
To describe, affirm, outline	Based on experience, analysis and consensus	Codes of ethics, standards of practice, regulations, instructions, guidebooks for interpreters
To perform the task	Based on experience, analysis and shared or individual experience	Teaching, research, stories, examples from "real life", etc.

and tacit knowledge related to interpreting. The function column tells what the knowledge is used for, the source column how it has emerged, and the contents column what is known and formulated in claims, values, stories, and so forth (cf. Rolf 1991: 27).

When the knowledge is verbalised, the function can be to describe, affirm or outline what is included in the knowledge. Codes of ethics, standards of practice and the various instructions and regulations for interpreters are most probably the tangible results of verbalised practical knowledge. Experienced interpreters have discussed and reached consensus on how to describe, affirm and outline good practice, and consequently the codes, standards, and so forth, have been written down.

Another function of verbalised knowledge can be to explain how to perform the task. The source of these verbalisations can be oral or written and may be based on either shared or individual experience or analysis. The tangible result of these verbalisations can be teaching and research publications or they can be stories and examples from real life told by interpreters. In a summary from a seminar for professional interpreters (Holmgren 2009: 19–20), the importance of stories and examples as ways of transferring practical and tacit knowledge was specifically emphasised.

4. A study of strategy selection in the interpreting process

The aim of this study of strategy selection is to learn which phases of the interpreting process informants at different stages of interpreting proficiency choose to give most weight to. The method used in this study is retrospection (Ericsson and Simon 1984; Krings 1986). The study examines how frequently the informants comment on problems they experienced or on their own acting at different phases of the interpreting process.

Three groups of informants participated in the study. These were new students taking their first interpreting course, advanced students with Interpreting Studies included as part of their Masters of Arts degree, and interpreters with 1–15 years of professional experience. When accepting interpreters with as little as one year of experience as informants, the decisive factor was that they should be professional interpreters, regularly receiving interpreting assignments and being accepted as interpreters by colleagues and clients (cf. Pilke 2004: 102).

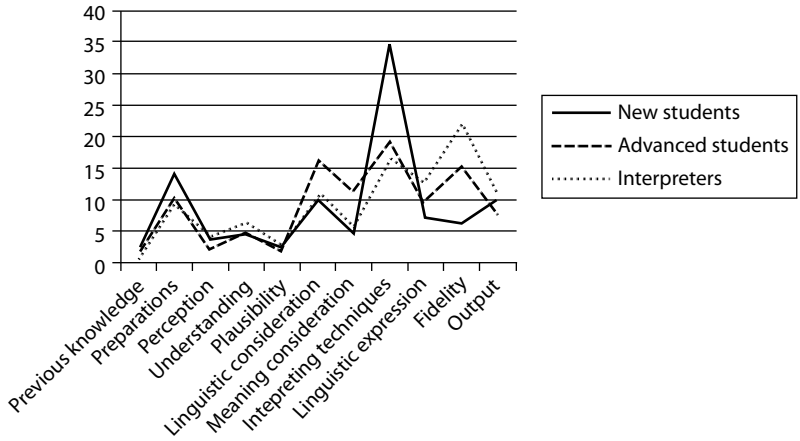


Figure 2. Allocation of retrospective oral comments from three groups of informants

Each group of informants consisted of 6–8 persons. They all carried out a simultaneous interpreting task from Finnish to Swedish. The source was a 7-minute speech originally given by a government minister at a conference, but for the purpose of the study performed as a video-taped simulation by another speaker. Both written and oral comments were obtained as research material (cf. Vik-Tuovinen 2006: 163). In this article, only results related to oral retrospective comments are discussed.

The informants listened to and commented on tape recordings of their own interpretation and their oral comments were tape-recorded, transcribed and analysed. The comments relate to different phases of the interpreting process. The distribution of the informants' comments is presented in Figure 2. The left-hand side of the graphic shows comments on preparations and input (previous knowledge, preparations, perception, understanding), while comments on transfer (plausibility, linguistic consideration, meaning consideration, interpreting technique) are included in the middle section and comments on the interpretation produced (use of linguistic expression, fidelity, output) are included on the right-hand side. The division of the process into phases is mainly based on Gile's sequential model of translation (Gile 1995: 102).

The volume of comments related to a certain phase of the process can reveal what aspects require attention and strategic measures, once the informant has reached a certain proficiency level. The average percentages of comments per informant group allocated to the different phases of the interpreting process are presented in Figure 2. The total number of comments received was 616, and in percentage terms, the distribution of comments was as follows: new students 18%, advanced students 36%, and interpreters 46%.

The new students were most concerned with preparing for the task and with interpreting techniques, where their main focus was on what they actually did or what problems they had with the interpreting techniques. The advanced students concentrated on linguistic considerations in the transfer phase and on fidelity in terms of

how well the content of their interpretation corresponded to the original, but to some extent they too focused on interpreting techniques. Techniques were also important for the interpreters, but their main emphasis was on fidelity. The advanced students made more comments on linguistic considerations than the other two groups. The likely reason is that for the new students, technical performance took up too much of their capacity and they did not have capacity left for linguistic considerations. The interpreters, on the other hand, did not need to concentrate on linguistic expressions in the transfer phase, as they had sufficient command of the languages and subject field involved.

A comparison of comments with the factors presented in Figure 1 (Section 2) shows that the new students focused on their own knowledge and mental capacity. The advanced students also paid attention to this, and additionally they also focused on linguistic options at the transfer phase and fidelity at the output phase, that is, on inter-textual norms. The advanced students seemed to know how to prepare for the task, they had at least to some extent mastered the simultaneous interpreting technique, and they were able to consider how to act in order to perform a good interpretation.

The comments of the interpreters and the advanced students relating to fidelity indicate that they are concerned with providing a reliable interpretation. A closer look at what the interpreters highlight in their comments supports the impression that despite having no audience other than the researcher during the task, they pay attention to potential clients and reflect upon the clients and their needs. They thus seem to have internalised their role as producers of an interpreting service and to emphasise situational factors as well as instrumental and intra-textual norms.

The way in which the advanced students commented on the interpreting task closely resembles the way in which the interpreters commented. The new students were different from advanced students in terms of both quantity and content of comments. The new students usually commented on problems without suggesting any strategies that could be used to solve the problem. The advanced students were able to reflect upon their problems in a more nuanced way, so in a manner more like the interpreters. This indicates that training increases students' ability to analyse their own performance and accordingly to develop their professional thinking and actions.

5. A questionnaire study on practical and tacit knowledge

The aim of the questionnaire study was to gain a general picture of the content, source and transfer of practical and tacit knowledge within the field of interpreting. The questionnaire study was based on five open-ended questions (Appendix 1). The questions include the term "practical knowledge", but the respondents were made aware that the term could be read as including practical skills and what is called tacit knowledge as well (cf. Section 3). Respondents were also encouraged to give real life examples to illustrate their answers. The questionnaire was answered by 11 advanced interpreting students and 10 interpreters with professional experience of 1–20 years.

The respondents were asked who they thought has practical knowledge of interpreting. Experienced interpreters ranked highest in the answers from both groups of respondents, being mentioned by 10 interpreters and 10 students. Other parties mentioned by several respondents in both groups were conference organisers, interpreting coordinators, interpreting teachers, interpreting students and clients. 5 of the interpreters also mentioned those in charge of technical equipment. These answers provide a basis for broadening the discussion about tacit knowledge, as the parties considered to possess practical and tacit knowledge include a variety of people involved in interpreting.

The respondents were also asked to explain what kind of practical knowledge interpreters need in their work. Most respondents from both groups, eight interpreters and nine students, mention knowledge needed for the current interpretation task, including preparation, knowledge about interpretation and required interpreting techniques. General knowledge and language skills are stressed by several respondents. Social skills and knowledge about practices and norms are mentioned by about half of the respondents from both groups. The students describe this kind of knowledge in a nuanced way, mentioning discretion, service-mindedness, interaction, confidentiality, crisis management, and so forth. This broad awareness among the students of the professional requirements indicates that practical knowledge has been an integral part of the training of these future interpreters.

Training and teachers are cited by all eleven students and seven of the interpreters as being the source of their own practical knowledge. Eight of the interpreters mention the importance of discussions with colleagues and observing colleagues, and two mention acquiring practical knowledge from a colleague before or during the assignment. Interpreting practice, discussion with other students, and observations of experienced interpreters are each mentioned by two student respondents. Written sources are mentioned by two interpreters and self-study by five students.

Several respondents reported that they have transferred practical knowledge themselves. The interpreters mentioned giving advice to colleagues, organisers, clients and speakers. Respondents from both groups mentioned that they had told family and friends about interpreting. Feedback included in training was mentioned by four of the students.

Both students and interpreters affirmed that to enhance the transfer of practical knowledge, interpreters should relay information about interpreting. They should participate as experts in the training of future interpreters and they should give information to organisers of events featuring interpretation and to clients. Secondly, the opportunity for students to participate as observers and to practice in real interpreting situations was mentioned by both groups of respondents. The third thing mentioned was making tacit and practical knowledge available to others. Networks and discussion forums on the web and teaching materials are suggested means of achieving that.

The conclusions from the study are that training is an important source of practical knowledge. The questionnaire also reveals the importance of communication

between different parties involved in interpreting, and the opportunities to gain knowledge by observations in real interpreting situations. Both groups of respondents emphasise that more ought to be written about tacit and practical knowledge, and they would like to see a wider range of organised forums and other channels for transferring knowledge within the field of interpreting.

6. An interview study on practical and tacit knowledge

The answers given by the respondents in the questionnaire study were quite short, and although the respondents were asked to give real life examples to illustrate their answers, very few examples were offered. In order to obtain more extensive descriptions of the phenomenon of practical/tacit knowledge, the questionnaire study was followed up by an interview study with four interpreters. Each of those respondents had experience of both simultaneous and consecutive interpreting in a variety of situations. The least experienced respondent had been working as an interpreter for a few years, the most experienced for several decades.

The interviews were recorded and transcribed. All the interviews were conducted in Swedish and the answers presented here have been translated into English. The respondents are called *A*, *B*, *C* and *D*. They have been working as interpreters for 3, 10, 20 and more than 35 years respectively, with *A* being the least and *D* the most experienced. The respondents were both male and female.

The respondents were told that they could talk about both practical and tacit knowledge in a broad sense. The interviews lasted about 25 minutes for each respondent and were based on questions given to them in advance of the interview (Appendix 2). We also encouraged them to raise other issues they felt were relevant.

In this presentation the focus will be on the respondents' descriptions of practical and tacit knowledge and their opinions on how the knowledge is related to situations and norms. These were two essential themes stemming from the interviews.

Interpreter *D* said that tacit knowledge is what you are not aware of but what you feel, and "you have to really think in order to become aware of it, aware of what you think was wrong in a situation or something, and then you start analysing yourself". *D* points out that tacit knowledge is problem-related, but also stresses that general knowledge and cultural knowledge are essential parts of an interpreter's tacit knowledge.

Interpreter *C* emphasised that tacit knowledge is based on experience; "it can be something a person who has been there many times before has learned over the course of time, so to say, it comes with experience". *C* also stressed that tacit knowledge can be based on knowledge about a particular assignment or client, and states "If you often work for the same client you learn how they want things".

The participants and the situation were also mentioned by interpreter *B*, who says of practical knowledge that "it is what you perhaps know about the speaker, or get to

know about the speaker and the situation, it has to do with all those present, the audience and the clients and the speaker, and the place and the kind of situation".

The least experienced, interpreter *A*, expressed the opinion that there are several things that should be included in tacit knowledge: gut feeling, intuition, emotional intelligence, and knowledge of human nature and patterns of different events. *A* illustrated the patterns with an example, stating that the interpreter should know how meetings move forward and "what happens when the discussion starts, what are the positions of the participating persons, and where do you think this discussion will lead". *A* expands on gut feeling, intuition, etc., by saying "much of it is what you cannot learn at school, but you either already have it or you may be able to learn it over time when you get more experience and learn more about life". Interpreter *A* concluded by saying: "These are things you learn when you […] have participated in different kinds of meetings […] you begin to understand the thing".

Issues arising from the interpreting situation are emphasised and explained in somewhat more detail in the interviews than in the questionnaire. Situational factors (cf. Figure 1, Section 2) seem to be relevant for practical and tacit knowledge. The two less experienced respondents emphasise aspects related to the interpreter's knowledge more often than the more experienced ones, but these aspects are not neglected by the more experienced respondents either.

The norms most strongly emphasised by the respondents are norms related to interpretation as received by a person who does not understand the source language and norms related to interpreting as a service for a client. The responsibility to pay attention to the participants – the intentions of speakers and the needs of listeners – are mentioned in some form by all the interpreters. *D* stresses that the interpreter has to help the speaker to get the message across just as it was intended. If something is emphasised by the speaker, then it must be emphasised in the interpretation as well, a merely neutral rendition is not adequate. According to *D*, that is what makes the difference, "if the interpreter succeeds in transferring the message the way the speaker delivers it, the interpreter's audience will feel that the interpreter was really good, even if the audience cannot explain this feeling". The responsibility to adapt to the wording of the speaker is also stressed by interpreter *D*, who says that it is important to use words that match the message of the speaker: *D* expands on the idea by stating "a left-wing speaker uses left-wing words".

According to all the respondents, the interpreter must adjust the interpretation to suit the listeners, taking the cultural background of the participants and geographic and socio-cultural variants of the language into account. All the respondents stress the benefits of recurring assignments for the same clients, and much of an interpreter's tacit knowledge is based on interviews associated with certain clients. As examples of issues related to the context, client or arrangement, *C* cites the manner in which contracts are drawn up, what kind of service the client wants, and what kind of situations the assignment may include. *C* also stresses the benefits of knowing what was discussed at the previous meeting and mentions that those are things that would

probably be passed on to a colleague who has not interpreted for that particular client before.

The descriptions given by the respondents suggest that practical and tacit knowledge can be related to problems or can be included in what an experienced interpreter generally knows about interpreting or comes to know about certain assignments or clients. Interpreters *A* and *B* mention that only when interpreters enter the profession they will understand the meaning of what they learned theoretically.

No clear-cut conclusions about differences between more experienced and less experienced interpreters can be drawn on the basis of only a few interviews. Generally, the descriptions given by the respondents highlight the role of the interpreter as a conscious part of the communication. The more experienced interpreters, *D* and *C*, seem to stress the interpreter's ability to use his practical/tacit knowledge and experience to "read the situation" in a somewhat stronger way, but the less experienced respondents also express their awareness of many central aspects related to situational factors. We might assume that less experienced interpreters are aware of essential situational aspects and more experienced interpreters have a better ability to act according to the situation, but more research would be needed before that assumption could be confirmed.

7. Discussion

The knowledge and skills required for performing good interpreting are quite extensive. Understanding of the many demands related to the situation, knowledge and norms involved in the interpreting process is important both to students of interpreting and practising interpreters. This knowledge should be learned and integrated into the performance of interpreting.

The study of selection of interpreting strategies presented in Section 4 indicates how knowledge and skills develop through training. While new students of interpreting concentrate on the interpreting technique, or how to manage to perform the interpretation, advanced students have learned the technique and can act more like professional interpreters. They pay attention to the quality of the interpretation in terms of proper linguistic expressions and fidelity. The thinking and acting of the up and coming interpreters has developed and some of the knowledge seems to be integrated in their performance.

The phenomenon of practical and tacit knowledge is discussed in Section 3. It is important to verbalise this kind of knowledge in order to develop awareness of it and to make it available for discussion with and transfer to others. The questionnaire study on practical and tacit knowledge presented in Section 5 shows that both interpreters and advanced interpreting students are aware of the diversity of the practical and tacit knowledge needed in interpreting. The questionnaire also reveals the importance of communication between different parties involved in the interpreted event and of

the opportunities provided by observations of real interpreting situations for eliciting and acquiring knowledge. The interview study presented in Section 6 confirms that interpreters stress the importance of the situational factors, and of their role as responsible communicators paying attention to the intentions of speakers and the needs of listeners.

The process of developing proficiency in any field involves several steps. Situations that first have to be analysed come to be recognised at once, knowledge is learned and integrated into procedures, procedures become more and more adapted to circumstances, and learners then understand the situation and their own roles as active and responsible parts of it (Bereiter and Scardamalia 1993: 77–120; Rantalaiho 1997: 247–250). The results of the studies presented in this article indicate a development of professional thinking and acting among interpreting students and interpreters following these steps.

More research needs to be undertaken on professional thinking and acting. A challenge for the whole field of interpreting is also to continue to outline and describe what good interpreting performance is, and to encourage people involved in the field to share their knowledge and experiences with newcomers, with colleagues and with other parties involved in interpreting.

References

Alexieva, B. 1997/2002. "A typology of interpreter-mediated events." In *The Interpreting Studies Reader*, F. Pöchhacker and M. Shlesinger (eds), 219–233. London/New York: Routledge.

Anderson, R. B. W. 1978. "Interpreter roles and interpretation situations: Cross-cutting typologies." In *Language Interpretation and Communication*, D. Gerver and H. W. Sinaiko (eds), 217–230. New York/London: Plenum Press.

Bartłomiejczyk, M. 2004. "Strategies of simultaneous interpreting from English into Polish and from Polish into English." (Doctoral dissertation. Faculty of Philology, University of Silesia, Poland.)

Bereiter, C. and Scardamalia, M. 1993. *Surpassing Ourselves*. Chicago/La Salle, Ill.: Open Court.

Chesterman, A. 1997. *Memes of Translation: The Spread of Ideas in Translation Theory*. Amsterdam/Philadelphia: John Benjamins.

Englund Dimitrova, B. 2005. *Expertise and Explication in the Translation Process*. Amsterdam/Philadelphia: John Benjamins.

Ericsson, K. A. and Simon, H. A. 1984. *Protocol Analysis: Verbal Reports as Data*. Cambridge, Mass.: MIT Press.

Garzone, G. 2002. "Quality and norms in interpretation." In *Interpreting in the 21st Century*, G. Garzone and M. Viezzi (eds), 107–119. Amsterdam/Philadelphia: John Benjamins.

Gerver, D. 1969/2002. "The effects of source language presentation rate on the performance of simultaneous conference interpreters." In *The Interpreting Studies Reader*, F. Pöchhacker and M. Shlesinger (eds), 52–66. London/New York: Routledge.

Gile, D. 1995. *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam/Philadelphia: John Benjamins.

Gustavsson, B. 2000. *Kunskapsfilosofi. Tre kunskapsformer i historisk belysning*. Stockholm: Wahlström & Widstrand.

Holmgren, K. 2009. *Tolkens tysta kunskap*. Härnösand: Härnösands folkhögskola.

Ivanova, A. V. 1999. "Discourse processing during simultaneous interpreting: An expertise approach." (Doctoral dissertation, University of Cambridge, UK.)

Johannesson, K. S. and Rolf, B. 1989. *Om tyst kunskap: Två artiklar*. Uppsala: University of Uppsala.

Jääskeläinen, R. and Tirkkonen-Condit, S. 1991. "Automatised processes in professional vs. non-professional translation: A think-aloud protocol study." In *Empirical Research in Translation and Intercultural Studies*, S. Tirkkonen-Condit (ed.), 89–109. Tübingen: Gunter Narr Verlag.

Kalina, S. 1998. *Strategische Prozesse beim Dolmetschen. Theroretische Grundlagen, empirische Fallstudien, didaktische Konsequenzen*. Tübingen: Gunter Narr Verlag.

Krings, H. P. 1986. *Was in den Köpfern von Übersetzern vorgeht*. Tübingen: Gunter Narr Verlag.

Lörscher, W. 1991. "Thinking-aloud as a method for collecting data on translation processes." In *Empirical Research in Translation and Intercultural Studies*, S. Tirkkonen-Condit (ed.), 67–77. Tübingen: Gunter Narr Verlag.

Mead, P. 2002. "Évolution des pauses dans l'apprentissage de l'interprétation consécutive." (Doctoral dissertation. Université Lumière Lyon 2, France.)

Namy, C. 1978. "Reflections on the training of simultaneous interpreters. A metalinguistic approach." In *Language Interpretation and Communication*, D. Gerver and H. W. Sinaiko (eds), 25–34. New York/London: Plenum Press.

Pilke, N. 2004. *Ordning och kaos: Om undervisning och inlärning vid tre seminarier på universitetsnivå* [Vaasan yliopiston julkaisuja. Tutkimuksia 263]. Vaasa: University of Vaasa.

Polanyi, M. 1966. *The Tacit Dimension*. Garden City, N.Y.: Doubleday.

Polanyi, M. and Prosch, H. 1975. *Meaning*. Chicago: University of Chicago Press.

Poyatos, F. 1987. "Non-verbal communication in simultaneous and consecutive interpretation: A theoretical model and new perspectives." *TexTconTexT* 2: 73–108.

Pöchhacker, F. 1994. *Simultandolmetschen als komplexes Handeln*. Tübingen: Gunter Narr Verlag.

Pöchhacker, F. and Shlesinger, M. (eds). 2002. *The Interpreting Studies Reader*. London/New York: Routledge.

Rantalaiho, K. 1997. "Osaamisen luonteesta." In *Muuttuva asiantuntijuus*, J. Kirjonen, P. Remes and A. Eteläpelto (eds), 246–252. Jyväskylä: Koulutuksen tutkimuslaitos/Jyväskylän ylipisto.

Rolf, B. 1991. *Profession, tradition och tyst kunskap. En studie i Michael Polanyis teori om den professionella kunskapens tysta dimension*. Övre Dalkarlshyttan: Bokförlaget Nya Doxa.

Shlesinger, M. 1989. "Simultaneous interpretation as a factor in effecting shifts in the position of texts on the oral-literate continuum." (Master's thesis, Faculty of Humanities, Tel Aviv University, Israel.)

———. 1997. "Quality in simultaneous interpreting." In *Conference Interpreting: Current Trends in Research*, Y. Gambier, D. Gile and C. Taylor, 123–132. Amsterdam/Philadelphia: John Benjamins.

Sunnari, M. 2004. "Kääntäminen ja tulkkaus monikielisyyden mahdollistajana." In *Kieli, teksti ja kääntäminen. Language, Text and Translation*, J. Tommola (ed.), 47–61. Turku: Turun yliopisto/Englannin kielen kääntäminen ja tulkkaus.

Tuomola, T. and Airila, A. 2007. "Hiljainen tieto yksilön ja työyhteisön näkökulmasta." In *Monikko – tasa-arvo monimuotoisissa työyhteisöissä*, K. Kauppinen, and J. Evans (eds), 72–83. Helsinki: Monikko-hanke.

Toury, G. 1995. *Descriptive Translation Studies and Beyond.* Amsterdam/Philadelphia: John Benjamins.

Vik-Tuovinen, G.-V. 2003. "Tulkki – erikoisalojen ja termien suurkuluttaja." *Terminfo* 3/2003: 7–9.

———. 2006. *Tolkning på olika nivåer av professionalitet* [Acta Wasaensia 153]. Vaasa: University of Vaasa.

———. 2008. "Oikeat toimeksiannot tulkkausopetuksessa." Paper presented at the KäTu Symposium, Turku 11–12.4.2008. Available at http://www.sktl.net/MikaEL/vol2/Vik-Tuovinen.pdf. (Accessed August 13, 2010.)

Wadensjö, C. 1992. "Interpreting as interaction: On dialogue interpreting in immigration hearings and medical encounters". (Doctoral dissertation, Linköping University, Sweden.)

Appendices

Appendix 1. Questionnaire

The questions in the questionnaire:

1. Who do you think has practical knowledge of interpreting?
2. What kind of practical knowledge do interpreters need in their job?
3. How and where have you acquired practical knowledge?
4. How have you yourself transferred practical knowledge to others?
5. How do you think the transfer and acquisition of practical knowledge can be enhanced?

Appendix 2. Interview

The questions in the interviews:

1. What do you think is included in an interpreter's practical knowledge or so-called tacit knowledge?
2. Is there something about clients, speakers, colleagues, organisers or other parties that you consider to be a very essential part of practical/tacit knowledge?
3. How can practical/tacit knowledge be learned?
4. What would you stress if you were to teach practical/tacit knowledge to a beginner or a less experienced colleague?
5. Have you yourself received or given advice about practical/tacit knowledge?

Results of the validation of the PACTE translation competence model

Translation problems and translation competence

PACTE Group

In alphabetical order: Allison Beeby, Mònica Fernández, Olivia Fox, Amparo Hurtado Albir, Anna Kuznik, Wilhelm Neunzig, Patricia Rodríguez, Lupe Romero, Stefanie Wimmer

Principal researcher: Amparo Hurtado Albir

Universitat Autònoma de Barcelona, Spain

The PACTE[1] Group has been carrying out experimental research into translation competence since 1977. The aim of this paper is to present the results obtained for one of our study variables: Identification and Solution of Translation Problems. A brief introduction to the PACTE Group's research project on translation competence is followed by a description of the variable Identification and Solution of Translation Problems, its indicators, and the instruments used for data collection. Finally, the results obtained are presented. These show that the translation problems identified by subjects varied greatly depending on the individual; directionality plays a role in the definition of the difficulty of translation problems; there was no relation between subjects' perception of the overall difficulty of a text to be translated and the acceptability of their solutions to translation problems; and, finally, the characterisation of translation problems would not appear to be a feature of translation competence.

Keywords: translation competence, translation problems, rich points, empirical-experimental research, acceptability of translations

1. In Spanish, the acronym PACTE corresponds to Process in the Acquisition of Translation Competence and Evaluation.

1. Introduction

The PACTE Group has been carrying out experimental research into translation competence and its acquisition in written translation since 1977. Research has been carried out from two complementary perspectives: (1) on the translation process, which involves gathering and analysing data obtained from experimental studies concerning the mental processes involved in translating and the competences and abilities required; and (2) on the translation product, which involves gathering and analysing data obtained from the results of the translation process (translated texts).

Our project is designed in two phases: (1) a first phase, now nearing completion, consisting of an empirical study of translation competence in which data concerning the knowledge and behaviour of professional translators is compared with that of foreign-language teachers with no experience in translation; (2) a second phase, begun in 2010, consisting of a longitudinal study of the acquisition of translation competence in trainee translators.

After first completing exploratory tests and a pilot study (PACTE 2002, 2005a, 2005b) to validate different aspects of our research design, an experiment was carried out to determine translation competence in 35 professional translators and 24 foreign-language teachers. Data have been collected on inverse and direct translations involving six language combinations: English-Spanish/Catalan, French-Spanish/Catalan, German-Spanish/Catalan.

The aim of this paper is to present the results obtained for the variable Identification and Solution of Translation Problems.

2. PACTE's research on translation competence

2.1 Theoretical model

The PACTE Group has defined translation competence as the underlying system of knowledge required to translate. In our opinion, translation competence (a) pertains to a field of expert knowledge, since it is not possessed by all bilinguals; (b) is predominantly procedural (operative) knowledge; (c) comprises different inter-related sub-competences; and (d) includes a strategic component which is of particular importance.

Following the PACTE model (cf. PACTE 2003), translation competence comprises five sub-competences as well as psycho-physiological components:[2]

2. PACTE (2003) believes that psycho-physiological aspects of translation competence should be differentiated from other sub-competences since they form an integral part of any expert knowledge. They are thus referred to as "components" and not as a sub-competence. This denomination is based on Bachman (1990).

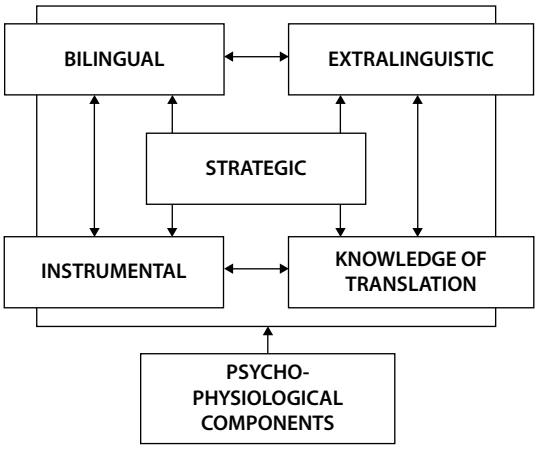


Figure 1. PACTE holistic model (PACTE 2003)

- *Bilingual sub-competence*: predominantly procedural knowledge required to communicate in two languages. It comprises pragmatic, socio-linguistic, textual, grammatical and lexical knowledge.
- *Extra-linguistic sub-competence:* predominantly declarative knowledge. It comprises general world knowledge, domain-specific knowledge, and bicultural and encyclopaedic knowledge.
- *Knowledge about translation*: predominantly declarative knowledge about translation and aspects of the profession. It comprises knowledge about how translation functions and knowledge about professional translation practice.
- *Instrumental sub-competence:* predominantly procedural knowledge related to the use of documentation resources and information and communication technologies applied to translation (dictionaries of all kinds, encyclopaedias, grammars, style books, parallel texts, electronic corpora, search engines, etc.).
- *Strategic sub-competence*: procedural knowledge to guarantee the efficiency of the translation process and solve problems encountered. This sub-competence serves to control the translation process. Its function is to plan the process and carry out the translation project (selecting the most appropriate method); evaluate the process and the partial results obtained in relation to the final purpose; activate the different sub-competences and compensate for any shortcomings; identify translation problems and apply procedures to solve them.
- *Psycho-physiological components*: different types of cognitive and attitudinal components and psycho-motor mechanisms, including cognitive components such as memory, perception, attention and emotion; attitudinal aspects such as intellectual curiosity, perseverance, rigour, the ability to think critically, etc.; and abilities such as creativity, logical reasoning, analysis and synthesis, etc.

Since all bilinguals possess knowledge of two languages as well as different degrees of extra-linguistic knowledge, the sub-competences that we believe are specific to translation competence are: strategic competence; instrumental competence and knowledge of translation. It is for this reason our research focuses on these three competences of which strategic competence is the most important since it interacts with all other sub-competences during the translation process, and serves to make decisions and solve problems.

2.2 Research design

Our general hypothesis is that the degree of expertise in translation (i.e. translation competence) is reflected in both the process and the product of translation.[3]

2.2.1 *Variables*

One independent variable and five dependent variables were selected for our study. The independent variable established was the degree of expertise in translation. In the independent variable, expertise in translation was manipulated in terms of greater or lesser degrees of expertise. Two categories were established: greater degree of expertise (translators with more than 5 years of professional experience guaranteed) and lesser degree of expertise (foreign-language teachers without any professional experience in translation and with more than 5 years' guaranteed experience as teachers).

Years of experience is obviously not the only criterion to define expertise. Additional criteria are required, such as acceptability of solutions to translation problems, knowledge of translation, decision-making behaviour, and so forth. The objective of our research is to identify the features that characterise expertise in translation (translation competence) and that serve to establish criteria and develop tests. For this reason we wish to observe the behaviour of the "best" translators from a number of different perspectives. For example, to determine the Efficacy of the Process (PACTE 2008), a comparison was made of the behaviour of translators that obtained the best and those that had obtained the worst results for the acceptability of their translations.[4]

The dependent variables in our study were (a) Knowledge About Translation; (b) Efficacy of the Translation Process; (c) Decision-making; (d) Translation Project; and (e) Identification and Solution of Translation Problems. During the experiment a further variable, Use of Instrumental Resources, was added. This was done because a large amount of data was collected on the use of instrumental resources by translators, and the indicators of the variables associated with the instrumental sub-competence

3. Our empirical and working hypotheses are based on the PACTE translation competence model (PACTE 2003).

4. See also 3.4 where the distribution in quartiles is used to study the relation between Coefficient of Satisfaction and Acceptability of Solutions.

(Decision-Making, and Identification and Solution of Translation Problems) were unable to provide a sufficiently detailed analysis of the data obtained.

When variables cannot be measured directly, indicators of the variables provide the necessary numerical data. In the PACTE experiment, a total of 18 indicators of the variables selected were identified, using data obtained in the exploratory tests and pilot study (see Table 1). Of these, the most relevant is the acceptability of subjects' translations, given that it reflects the quality of their translations (an important aspect of their translation competence). The indicator Acceptability is used as a transversal indicator in conjunction with indicators of all the variables under study in order to determine the relationship that exists between the results obtained in these indicators and the quality of subjects' translations.

2.2.2 *Sample and experimental tasks*

Two experimental groups were selected from the population of professionals working with foreign languages: (i) professional translators with guaranteed experience in translation (potentially with a greater degree of expertise in translation); and (ii) foreign-language teachers with no professional experience in translation (potentially with a lesser degree of expertise in translation). All subjects were required to be native speakers of Spanish and/or Catalan and to work in a professional capacity with German, French or English as their foreign language, and the period of their professional activity should be equivalent.

Given that professional translators are not necessarily experts,[5] and also for the economy of the experiment (Giegler 1994), other selection criteria were years of professional experience as translators and translation as their main source of income. A further criterion was that the translators should not be specialised in any particular field of translation. To ensure the absence of confounding variables, it was decided that translators should not be specialists in any particular field of translation, since specialisation in any specific field – literary, legal, audiovisual, etc. – could distort results.[6]

A questionnaire was used to select subjects for both groups who fulfilled these criteria.

Thirty-five professional translators[7] and twenty-four foreign-language teachers participated in the experiment. The translators included in the study had an average of 7.51 years of experience in translating, the average percentage of their income from translating was 86.43%, and their experience included translating a wide range

5. In Spain, officially recognised expertise, i.e. accreditation of translators (sworn translators), is limited to legal translation.

6. This was done to ensure the homogeneity of the sample. In future, studies should be carried out on translation competence in the fields of literary translation, legal translation, audiovisual translation, etc.

7. Translators were contacted through professional associations.

of texts into their native language. Foreign-language teachers all had a minimum of five years' experience of teaching in the Spanish Ministry of Education's Modern Language School (*Escuela Oficial de Idiomas*). Subjects were paid for their translations, simulating a real-life translation task.

Subjects performed the following tasks: (1) direct translation; (2) completion of a questionnaire about the problems encountered in the translation; (3) inverse translation; (4) completion of a questionnaire about the problems encountered in the translation; (5) completion of a questionnaire about translation knowledge; (6) participation in a retrospective interview.

Each of these tasks provided data for analysis. Further data were obtained from real-time recordings of subjects' actions during the translation process using the software programs PROXY and Camtasia,[8] and direct observation.

Table 1 summarises the most important information relating to the research design.

The methods used and the findings obtained for the four variables Knowledge About Translation, Efficacy of the Process, Decision-making, and Acceptability have been published in PACTE (2007a, 2007b, 2008). The results obtained for the variable Translation Project and their triangulation with those obtained for the variable Knowledge About Translation has been published in PACTE (2011).

2.2.3 *Data analysis*

Data obtained for each study variable was analysed and the results triangulated, thereby (1) relating results for the different indicators of study variables; (2) relating translators' and teachers' behaviour; (3) relating translators' and teachers' performance in direct and inverse translation; and (4) relating the results for indicators of all variables and the indicator Acceptability.

For the purposes of data collection and analysis, the decision was made to focus attention on specific source-text segments that contained "prototypical" translation problems, i.e. the most salient, characteristic and difficult problems in a text. These we refer to as Rich Points. The decision to focus data collection on selected Rich Points was taken to facilitate data collection, following Giegler's concept of "scientific economy" (Giegler 1994), and the triangulation of data (cf. PACTE 2007b, 2008, 2009). The Rich Points selected were determined as a result of exploratory studies and pilot tests carried out prior to the final experiment (PACTE 2002, 2005a, 2005b). The texts selected for use in the experiment, together with five Rich Points identified in each, were trialled in the pilot study carried out in 2004 (PACTE 2005a, 2005b). The Spanish source text used for inverse translation and the English source text used for direct translation are included in Appendix 1.

---

8. PROXY is a program (compatible with Windows) designed for the remote control of computers and users connected to a network. Camtasia records the subject's actions on the computer in real time and stores these recordings for subsequent study and data collection.

Table 1. Relevant aspects relating to the research design (adapted from PACTE 2005a, 2005b)

KNOWLEDGE ABOUT TRANSLATION (related to the knowledge about translation sub-competence)	
Conceptual Definition	The subject's implicit knowledge about the principles of translation and aspects of the translation profession.
Indicators	Dynamic index and coherence coefficient.
Instruments	Questionnaire on knowledge about translation.
EFFICACY OF THE TRANSLATION PROCESS (related to the strategic sub-competence)	
Conceptual Definition	Optimum relationship between time taken to complete a translation task and the acceptability of the solution.
Indicators	Total time taken; time taken at each stage of the translation process (orientation, development, revision; based on Jakobsen 2002); acceptability.
Instruments	Translations, direct observation chart, PROXY and Camtasia recordings.
DECISION-MAKING (related to strategic and instrumental sub-competences)	
Conceptual Definition	Decisions made during the translation process which involve the use of automatised and non-automatised cognitive resources (internal support) and the use of different sources of documentation (external support) (Alves 1995, 1997).
Indicators	Sequences of actions; acceptability.
Instruments	Translations, direct observation charts, PROXY and Camtasia recordings.
TRANSLATION PROJECT (related to the strategic sub-competence)	
Conceptual Definition	The subject's approach to the translation of a specific text and of the units it comprises.
Indicators	Dynamic index in the overall translation project and that of each Rich Point; coherence between the overall translation project and that of each Rich Point; acceptability.
Instruments	Translation problems questionnaire and retrospective interview.
IDENTIFICATION AND SOLUTION OF TRANSLATION PROBLEMS (related to the strategic sub-competence and the knowledge about translation sub-competence)	
Conceptual Definition	Difficulties encountered by the subjects when carrying out a translation task.
Indicators	Coefficient of perception of the overall difficulty of the text; subjects' identification of Rich Points; characterisation of Rich Points identified by subjects; coefficient of subjects' satisfaction with the solution found for each Rich Point; type of internal support used to solve each Rich Point; acceptability.
Instruments	Translation problems questionnaire and retrospective interview.
USE OF INSTRUMENTAL RESOURCES (related to the instrumental competence)	
Conceptual Definition	Strategies used when consulting documentary resources in electronic format (websites, dictionaries and encyclopaedias in CD-ROM).
Indicators	Variety of resources, number of searches, time spent on searches (total and for each phase).
Instruments	PROXY/Camtasia recordings, Catalogue of searches.

The use of Acceptability as a transversal indicator when analysing study variables should be noted. Acceptability is an indicator of the quality of the translation product. Quantitative and qualitative analysis of the data collected in our exploratory tests (PACTE 2002) and pilot study (PACTE 2005a, 2005b) confirmed the importance of this indicator in measuring subjects' expertise in translation. It is the only indicator that is used in conjunction with specific indicators of each variable (see Table 1). In our research project, Acceptability is defined in terms of whether or not the solution effectively communicates (a) the meaning of the source text; (b) the function of the translation (within the context of the translation brief, the readers' expectations, genre conventions in the target culture); and (c) makes use of appropriate language.

The following criteria were used to identify acceptable, semi-acceptable and not acceptable solutions (PACTE 2008):

- *Acceptable solution* (A): The solution activates all the relevant connotations of the ST in the context of the translation related to the meaning of the ST, function of the translation and language use.
- *Semi-acceptable solution* (SA): The solution activates some of the relevant connotations of the ST and maintains the coherence of the TL in the context of the translation related to the meaning of the ST, function of the translation and language use.
- *Not acceptable solution* (NA): The solution activates none of the relevant connotations of the ST or introduces connotations that are incoherent in the context of the translation related to the meaning of the ST, function of the translation and/or language use.

Our definition of what constituted acceptable, semi-acceptable and non-acceptable solutions for each Rich Point was based on these criteria, taking into account the features of the Rich Point that had previously been established. The evaluation of solutions was carried out using the procedure of peer review for each Rich Point in each language combination. To ensure reliability in the application of the criteria established in all language combinations the group as a whole was consulted when any doubts arose.

Thus, each translation solution given was classified as A, SA, NA. Twenty-seven possible permutations were obtained (PACTE 2007b, 2008). Subsequently numeric values were then assigned to each category: A = 1; SA = 0.5; NA = 0 (Table 2). Finally, a descriptive analysis was made of the numerical values obtained for each RP, in both direct and inverse translation, and for each experimental group.

Table 2. Acceptability scores across groups (PACTE 2008, 2009)

Acceptability		**Translators**	**Teachers**
Direct translation	Mean	.73	.49
	Median	.80	.45
Inverse translation	Mean	.52	.48
	Median	.50	.40

Acceptability scores for direct and inverse translation (PACTE 2008, 2009) showed that the group of translators obtained more acceptable results in their translations than the group of foreign-language teachers. Results also showed that the difference in the acceptability of the results obtained in both groups is much greater in direct translation (see Table 2).

3. Identification and solution of translation problems

According to Wilss, there is no uniform concept in Translation Studies of what constitutes a translation problem nor any appropriate, reliable conceptual framework (Wilss 1996: 47). Lörscher draws attention to the fact that the debate over translation problems has been largely speculative and that little in the way of empirical study has been carried out (Lörscher 1991: 12).

3.1 Translation problems

Until the 1980s, translation problems were approached almost exclusively from a linguistic point of view (Vinay and Darbelnet 1958; Mounin 1963; Catford 1965; Vázquez Ayora 1977; Scavée and Intravaia 1979; Hönig and Kussmaul 1982; etc.). Since then, other authors have taken a wider view and have identified a number of different types of problems, e.g. textual, pragmatic, cultural and linguistic (Nord 1991); textual, pragmatic and semiotic (Presas 1996); and text-linguistic, extralinguistic, pragmatic and instrumental (Hurtado 2001).

From a cognitive perspective, translation problems have been linked to the cognitive processes taking place while translating, and to the concept of translation competence. Krings (1986) used the notion of translation problems as a category in his analysis of subjects' comments recorded in think-aloud protocols. He claims that the presence of a problem is evidenced implicitly or explicitly by subjects during the translation process through pauses, use of strategies, omissions, corrections, and so forth. Krings differentiates between two types of indicators: primary (explicit reference made by the subject; recourse to dictionaries; omissions in translation) and secondary (tentative translations; corrections; notes made on the source text). He suggests that translation problems may be divided into problems of reception, problems of production, and problems of reception-production (in which both types of problems are combined).

Bell (1998) also notes that translation problems are part of the process of transfer of a text (reception, production) and are associated with non-automatic processes: "A translation problem is some part of the process of transfer, whether deriving from the reception of the source text or the production of the target text, which makes analysis or synthesis non-automatic" (Bell 1998: 188). Citing Lörscher, he goes on to define the role of strategies in solving translation problems: "On the basis of this definition,

a translation strategy is a potentially conscious procedure for solving a problem faced in translating a text, or any segment of it" (Lörscher 1991: 76; cit. Bell 1998: 188).

Kiraly (1995: 99–105) believes that there is no clear dichotomy between controlled and uncontrolled processes in a translator's mind. Thus, in his model of the cognitive processes involved in translation, he suggests that translators have a relatively uncontrolled processing centre (intuitive, less conscious) and a relatively controlled processing centre (strategic, more conscious). Like Bell, Kiraly associates translation problems with non-automatic processes. Translation problems emerge from the intuitive workspace when automatic processing does not produce tentative translation elements. These problems are considered in the controlled processing centre and a strategy is chosen and implemented in an attempt to deal with them (1995: 105).

Another subject of debate in Translation Studies is whether or not all problems translators have when translating may be deemed translation problems. Nord distinguishes between problems and difficulties in translation. Thus, a translation problem is "an objective problem which every translator (irrespective of his level of competence and of the technical conditions of his work) has to solve during a particular translation task" (1991: 151). Difficulties in translation are "subjective and have to do with the translator himself and his specific working conditions" (1991: 151).

3.2 Translation problems in PACTE's research on translation competence

We believe that one of the main characteristics of translation competence, like all expert knowledge, is the ability to solve problems. The solution of translation problems involves different cognitive operations within the translation process, and requires constant decision-making on the part of the translator.

In PACTE's holistic model of translation competence, this ability to solve problems is directly linked to the strategic sub-competence. Strategic competence is the most important of all the sub-competences since it serves to control the translation process and is involved in decision-making processes. In order to obtain data concerning the role of strategic competence in our study of the translation process, the following variables were selected: Efficacy of the Process; Decision-making; Translation Project; and Identification and Solution of Translation Problems (PACTE 2008).

3.2.1 *Prototypical translation problems*

We believe that being competent in a profession implies being able to effectively carry out the tasks required of that profession, and therefore, of being able to solve the types of problems most frequently encountered in the course of one's professional activity.

When selecting the texts used in our experiment, therefore, an effort was made to find texts that posed the types of translation problems most frequently encountered by professional translators. As mentioned above (Section 2.2.3), the decision was made to focus data collection and analysis on specific source-text segments that contained this type of translation problem, i.e. those we refer to as Rich Points.

Table 3. Categories of translation problems

Linguistic problems	Lexis (non-specialised) and morphosyntax. Problems of both comprehension and re-expression may be involved.
Textual problems	Coherence, cohesion, text type and genre, and style. Problems of comprehension or re-expression may be involved, and are associated with differences in the way texts function in each language.
Extralinguistic problems	Cultural, encyclopaedic and subject-domain knowledge. Problems that also derive from cultural differences.
Problems of intentionality	Difficulty in understanding information in the source text (intertextuality, speech acts, presuppositions, implicatures). Problems of comprehension.
Problems relating to the translation brief and/or the target-text reader	Difficulties (affecting reformulation) which, from a functionalist point of view, would affect all Rich Points.

When identifying the Rich Points in the SL texts selected – one involving direct and the other involving inverse translation[9] – five categories of translation problems were considered (see Table 3).

Our research has shown that translation problems are not uni-dimensional – they may share features of several different categories.

In the text used for inverse translation (see Appendix 1) the following Rich Points were selected:

RP1: *indiano … fortuna del Americano* – extralinguistic, textual problem
RP2: *gobierno alfonsino* – extralinguistic problem
RP3: *desenfreno y dilapidación* – linguistic problem of re-expression
RP4: *la geografía comarcal de Cataluña* – problem of intentionality
RP5: *común … trona* – extralinguistic, textual, and problem of intentionality.

For the purposes of direct translation, parallel texts in English, French and German on the subject of computer viruses were used (see Appendix 1). The English source text (EST) was "E-mail virus strikes in new form" (*The Guardian*, June 6, 2003);[10] the German source text (GST) was "Wurm in der Leitung" (*Frankfurter Allgemeine Zeitung*, June 14, 2003); and the French source text (FST) was "Bugbear.B, le virus informatique qui lit par-dessus l'épaule de ses victimes" (*Le Monde*, June 13, 2003).

In the texts used for direct translation, the types of problems selected were:

9. As an example of the texts used, Appendix 1 presents the English source text for direct translation and Spanish source text for inverse translation.

10. We would like to thank F. Alves for providing this text.

RP1 Title: Problem of intentionality; textual problem.

RP2 Technical term: *keylogger* (EST); *Download-Verzeichnis* (GST); *Édition de logiciels antivirus* (FST). Linguistic problem of re-expression; extralinguistic problem.

RP3 Reference: *propagation rate … doubled … surge* (EST); *Schädling … E-Mail Würmer … Vorgängervariante* (GST); *Le ver … résurgence … ses congénères* (FST). Textual problem.

RP4 Element in apposition: *"Trojan horse" program which could allow a hacker to take remote control of infected machines* (EST); *Dateien-Tauchbörse Kazaa* (GST); *Soumissions, des communications du virus* (FST). Textual problem; problem of intentionality.

RP5 Element involving difficulties in comprehension and reformulation: *Cheltenham-based virus filtering firm* (EST); *Tastatureingaben von PC-Nutzern nach Kreditkartennummern und Ähnlichem überwacht* (GST); *Enregistrer les caractères tapés sur le clavier* (FST). Problem of intentionality; linguistic problem of re-expression.

It was assumed that, from a functional concept of translation, the purpose of the text and the target reader affected all Rich Points.

3.2.2 *Translation problems and cognitive processes*

From a cognitive point of view, and based on Bell (1998) and Kiraly (1995), we believe that a translation problem exists when "automatised"[11] solutions, i.e. spontaneous and immediate solutions, are not found for source-text segments in translation and different strategies are then put into effect to solve them. These strategies we believe are actions that are used to obtain specific objectives or outcomes (Pozo, Gonzalo and Postigo 1993).

Following Alves (1995, 1997), we believe that translators use two types of resources during the translation process: internal support (automatised and non-automatised cognitive resources) and external support (all kinds of information resources).

Thus, in our study of the variable Decision-making (PACTE 2005b, 2009) different sequences of actions (strategies) were found to be used to solve translation problems:

– *Internal Support:* the Definitive Solution is based exclusively on internal support, with no consultation prior to the Definitive Solution being adopted. Example: Definitive Solution; Provisional Solution → Definitive Solution.

11. We prefer the use of the term "automatised" and "non-automatised" rather than "automatic" and "non-automatic", since we believe that when translating a text, intentionality is always involved and therefore, as Kiraly (1995) observes, there are no totally unconscious processes involved in translation. It would be more appropriate to speak of "automatised" processes, the product of experience.

- *Predominantly Internal Support:* the Definitive Solution is based essentially on internal support, i.e. any combination of consultations which does not include a consultation of bilingual resources from which the variant offered is adopted in translation. Example: Provisional Solution → Consultation of Alternative Resources → Consultation of Alternative Resources → Consultation of Bilingual Resources (but the variant is not adopted in the translation) → Definitive Solution.
- *Predominantly External Support:* the Definitive Solution is based essentially on external support, i.e. any combination of consultations that includes consultations of bilingual resources from which the variant offered is adopted in translation. Example: Provisional Solution → Consultation of Alternative Resources → Consultation of Alternative Resources → Consultation of Bilingual Resources (and the variant offered is adopted in the translation) → Definitive Solution.
- *External Support:* the Definitive Solution is based exclusively on consultation of bilingual resources, from which the variant offered is accepted in the translation. Example: Consultation in Bilingual Resources (and the variant offered adopted in translation) → Definitive Solution.

Regarding the use of Internal Support, following Bell and Kiraly, we believe that automatised internal support is used when a translator finds no problem in translating a source-text segment. When a problem *is* encountered, non-automatised internal support, i.e. cognitive strategies, is used to solve the problem.

3.3 The variable "Identification and Solution of Translation Problems"

The variable was designed to obtain data on the way in which the process of identification and resolution of translation problems evolves and its relation to translation competence. The variable Identification and Solution of Translation Problems is defined as difficulties encountered by subjects when carrying out a translation task. When determining this variable the following questions were asked:

- *Is the ability to identify and solve translation problems really a characteristic of translation competence? Do translators and non-translators behave in the same way? Do they perceive the same degree of difficulty in translating a text? Do they identify the same kinds of problem? Do they describe the problems in the same terms? Is there any difference between direct and inverse translation in relation to translation problems?*
- *Is there any link between translation problems and non-automatised cognitive processes? Do translators and non-translators make the same use of automatised and non-automatised internal support?*
- *Can we distinguish between translation problems (objective) and difficulties in translation (subjective)?*
- *Is there a link between the difficulties encountered by subjects and the acceptability of results?*

– *Is the ability to evaluate the solutions of translation problems a characteristic of translation competence?*

The desire to obtain answers to these questions informed the selection of indicators for the variable under study both in relation to the text as a whole as well as the selection of prototypical translation problems (Rich Points). These indicators were:

1. Coefficient of perception of the overall difficulty of the translation of the text.
2. Subjects' identification of prototypical translation problems.
3. Characterisation of the prototypical translation problems identified by subjects.
4. Coefficient of subjects' satisfaction with their solution for each of the prototypical translation problems identified.
5. Type of internal support used to solve prototypical translation problems.

Data for the variable were collected using a translation problems questionnaire and a retrospective interview, both of which focused on the Rich Points selected in each of the source texts to be translated (see Appendix 2). Questions were designed to provide data relating to the indicators established, specifically:

– *How difficult do you think this text is to translate?* (Indicator 1)
– *What were the main problems you found when translating this text? Name 5 and answer the following questions about each.* (Indicator 2)
– *Why was it a problem?* (Indicator 3)
– *Are you satisfied with the solution?* (Indicator 4)

Information concerning the identification and characterisation of the Rich Points which was not obtained from subjects' answers in the questionnaire was subsequently solicited in the retrospective interview. The information obtained from the interview was then used to complement the information obtained from the questionnaire and the data as a whole was analysed.

Data concerning the type of internal support (either non-automatised or automatised, see Section 3.2.2) was obtained by cross-referencing the data for the use of internal support in the variable Decision-making (PACTE 2009) with subjects' identification of problems in the questionnaire and retrospective interview. If a problem had existed, it was deemed that non-automatised internal support was used (thinking), and if it had not existed, it was deemed that automatised internal support was used. These data were then triangulated with the acceptability of the solutions obtained.

This article presents only the results obtained for the first four indicators of the variable Identification and Solution of Translation Problems. The fifth indicator (type of internal support), related to the indicator Sequences of Actions and the variable Decision-making (PACTE 2009), requires more in-depth explanation than this article permits and will be the subject of a future publication.

4. Results

4.1 Coefficient of perception of the overall difficulty of the text

Subjects' perception of the overall difficulty of the texts to be translated was recorded on a scale of 1–30. For the purposes of data analysis the scores recorded were converted to 100 to obtain a coefficient of 0–1. Scores were classified by language to determine whether or not the degree of difficulty perceived was in any way influenced by language.

4.1.1 *Results obtained for direct translation*

The average of the coefficients obtained for translators and foreign-language teachers in direct translation were as shown in Table 4.

Results show that the group of translators in general perceived the text to be relatively easy to translate.

A comparison of both the teachers' and translators' perception of the overall difficulty of the text shows that whilst there is little difference between the difficulty perceived by the French and German translators (.30) and the French- and German-language teachers (.34; .38), the English-language teachers (.56) perceived the text to be much more difficult than the English translators (.27). As a result of the marked difficulty perceived by teachers of English, the difference between translators' and teachers' perception of the overall difficulty of the translation of the text is relevant (translators: .28; teachers: .43). The group of teachers perceived the text to be more difficult.

The homogeneity of the results obtained for the group of translators (English, French and German) for the coefficient of perception of the overall difficulty of the text contrasts with that of the group of teachers.

Table 4. Mean coefficient of perception of the overall difficulty of the text (direct translation)

	Translators	**Teachers**
English	.27	.56
French	.30	.34
German	.30	.38
Overall	**.28**	**.43**

4.1.2 *Results obtained for inverse translation*

The average of the coefficients obtained for translators and foreign-language teachers in inverse translation were as shown in Table 5.

These results show that subjects in both groups perceived the text to be difficult. However, whilst the degree of difficulty perceived by the group of translators was similar (.66; .60; .64), it varied within the group of teachers: teachers of English and

Table 5. Mean coefficient of perception of the overall difficulty of the text (inverse translation)

	Translators	Teachers
English	.66	.77
French	.60	.56
German	.64	.77
Overall	**.63**	**.70**

German perceived the text to be more difficult to translate (.77) than teachers of French did (.56). The proximity of the French and Spanish languages may well account for the difference in difficulty perceived by teachers of French. As a result of the low level of difficulty of the text perceived by teachers of French, the difference between translators' and teachers' perception of the overall difficulty of the translation of the text is not relevant (translators: .63; teachers: .70).

The homogeneity of the results obtained for the group of translators (English, French and German) for the coefficient of perception of the overall difficulty of the text in inverse translation, as in direct translation, contrasts with that of the group of teachers and would indicate that translators were not influenced by the language into which they had to translate.

A comparison of the results obtained for direct and inverse translation shows that both groups find inverse translation (translators: .63; teachers: .70) much more difficult than direct translation (translators: .28; teachers: .43).

4.1.3 *Coefficient of perception of the difficulty of the text and acceptability*

When subjects are ranked according to the level of acceptability of their solutions to translation problems, no tendency or pattern can be observed between acceptability and the overall difficulty of the text perceived by translators and teachers in inverse or direct translation.

Pearson's coefficient of correlation (Pearson r) shows that the correlation between acceptability and the perceived difficulty of the text is not statistically significant in either of the two experimental groups (cf. Table 6). No tendencies can be observed in inverse or direct translation. The perceived difficulty of the text does not therefore, lead to more acceptable or less acceptable solutions.

Table 6. Coefficient of perception of the difficulty of the text and acceptability (significance level for Pearson r for translators > 0.35; teachers > 0.39)

	Pearson (r) coefficient of correlation	Degree of freedom	Significance
Translators (direct)	0.13	32	Not significant
Teachers (direct)	0.01	22	Not significant
Translators (inverse)	0.04	32	Not significant
Teachers (inverse)	0.19	22	Not significant

4.2 Subjects' identification of prototypical translation problems

A calculation was made of the percentage of Rich Points (RP) identified by each group (translators and teachers) in both the questionnaire and the retrospective interview.

In direct translation (cf. Table 7), more translators (62.9%) than teachers (33.3%) identified the title (RP1: problem of intentionality) as a translation problem. Fewer translators than teachers identified RP3 and RP4 (textual) as problems – these are precisely the Rich Points that translators would have solved most acceptably (0.8 in each case; PACTE 2011). The problem least often identified by both translators (22.9%) and teachers (33.3%) was RP5: problem of intentionality and linguistic problem of re-expression).

Since all five Rich Points posed problems for subjects in both groups, these results would confirm that the Rich Points selected for our experiment were appropriate.

In inverse translation (cf. Table 8), many more teachers (70.8%) than translators (57.1%) identified RP3 *desenfreno y dilapidación* (linguistic problem of re-expression) as a problem. The percentage of subjects who identified the remaining Rich Points as translation problems was similar.

It should be noted that the percentage of Rich Points identified as translation problems was higher in inverse translation than in direct translation. The greater overall difficulty of the text perceived by subjects in inverse translation is thus reflected in the high percentage of Rich Points identified as problems.

The results confirm that the Rich Points selected for our experiment constituted an effective sample of problems found in inverse translation since they were identified by most of the subjects.

Finally, each subject selected different prototypical problems. Subjective difficulties, depending on personal knowledge and skills, may well account for the difference in which of the Rich Points was identified as a problem and by whom.

Table 7. Identification of prototypical translation problems (direct translation)

	RP1 Title	**RP2 Technical term**	**RP3 Reference**	**RP4 Apposition**	**RP5 Comprehension and reformulation**
Translators	62.9%	51.4%	54.3%	40.0%	22.9%
Teachers	33.3%	45.8%	62.5%	50.0%	33.3%

Table 8. Identification of prototypical translation problems (inverse translation)

	RP1 ***indiano … fortuna***	**RP2** ***gobierno alfonsino***	**RP3** ***desenfreno y dilapidación***	**RP4** ***geografía comarcal***	**RP5** ***Común … trona***
Translators	71.4%	65.7%	57.1%	68.6%	68.6%
Teachers	66.7%	66.7%	70.8 %	62.5%	75.0 %

4.3 Characterisation of the prototypical translation problems identified

An analysis was made of subjects' characterisation of the Rich Points identified as problems in the questionnaire and retrospective interview (Question: *Why was it a problem?*). The range and lack of precision of subjects' answers to this open-ended question made it necessary at times to interpret their answers. It was observed that it was difficult for subjects to explain the difficulties they encountered given the procedural cognitive processes involved in translation and the automisation of all expert knowledge.

Subjects' answers were assigned to the categories below using the following criteria:

- *Linguistic difficulties*: when subjects refer to lexical or morphosyntactical difficulties. A distinction is made between linguistic difficulties of comprehension (C) and re-expression (R), in order to determine whether a specific difficulty is due to shortcomings in a subject's knowledge of the source-text or target-text language, and whether or not it occurs in the same way in direct or inverse translation in each of the subject groups.12 In some cases both types of linguistic difficulties (C and R) arose. When dealing with technical terms, reference is made to a linguistic difficulty of re-expression of terminology (T) since this is a special case linked to extralinguistic knowledge. Examples of how answers were assigned: "I've never heard this word before" (linguistic C); "I didn't know how to say it in X" (linguistic R); "Difficulty in finding the exact term" (linguistic R T).
- *Textual difficulty*: when subjects refer to difficulties of coherence, cohesion, text types (genre conventions) and style. In this case, no difference is made between comprehension and re-expression since this type of difficulty is considered to be one of contrastivity between two languages. Example: "I've already mentioned the American"; "Because I wanted it to conform to the characteristics of the genre".
- *Extralinguistic difficulty*: when subjects refer to difficulties relating to a specific field of knowledge (specialised concepts) or bicultural or encyclopaedic knowledge. Examples: "The concept does not exist in German"; "An equivalent territorial and administrative division does not exist". With regard to terminology (specialised concepts), when reference is made to difficulties of understanding concepts expressed in the source text (example: "I didn't know the meaning of the word") these are assigned to the category of extralinguistic difficulties. In those cases in which reference is made to difficulties of finding the word in the target text language, terminology is classified as Linguistic R T. Example: "I had difficulty in finding the exact term".

---

12. We consider this to be of importance when determining the approriateness of our sample: if marked shortcomings had been observed in subjects' knowledge of the source language in either groups, acceptability would have be conditioned by those shortcomings and not by other features of translation competence.

- *Difficulties of intentionality*: when subjects refer to difficulties in understanding the information contained in the source text. A clear distinction is made between difficulties related to intentionality and difficulties of comprehension that are purely linguistic in nature. When the difficulties referred to by subjects are those that involve the information load of a Rich Point, they are classified as difficulties of intentionality. When they refer to lexical or syntactical difficulties, they are considered to be linguistic. Example: "I didn't know what it meant in this context" (intentionality); "I didn't know what it meant" (linguistic C – when referring to non-specialist terms).
- *Difficulties relating to the function of the text and the target reader*: when subjects refer to problems relating to the function of the translated text (according to the translation brief) and/or the target reader. Example: "Because it depends on the reader's knowledge".

4.3.1 *Results in direct and inverse translation*

The percentage of each of the categories used to describe the Rich Points identified was then calculated for each group (translators and teachers). It was, however, difficult to observe any relevant differences between the two groups as a result of quantitative analysis for two reasons: (i) subjects' answers were often confusing and therefore difficult to categorise; (ii) the number of subjects in each case was small (only those who had identified the problem and had been able to attribute it to a specific category of problem).

Taken overall, there would appear to be no notable difference between the way in which translators and teachers characterise the translation problems they identified. The only differences observed were the following:

- There was a greater tendency on the part of teachers to describe difficulties as linguistic, e.g. in direct translation they described RP1 (title) and RP4 (element in apposition) as linguistic difficulties of expression and RP5 (*común … trona*) as a linguistic difficulty of comprehension.
- There was a difference in the way in which problems of intentionality were described (in direct translation: RP1 the title; in inverse translation: RP4 *geografía comarcal* and RP5 *común … trona*). Most teachers described the difficulties posed by these Rich Points as linguistic, whilst most translators assign them to a wider range of categories (function of the text, textual difficulties and difficulties of intentionality), underlining the multidimensionality of translation problems.

On the basis of these results, we cannot affirm that the ability to explain the nature of a translation problem is a characteristic of translation competence.

4.4 Coefficient of subjects' satisfaction

To calculate this coefficient, the answers to the questionnaire (YES/NO/Partially) were given numerical values: YES = 1; NO = 0, P = 0.5. A coefficient of 0–1 was calculated for each subject.

The following coefficients were calculated: (1) coefficient of subjects' satisfaction; (2) coefficient of subjects' satisfaction for each Rich Point; (3) coefficient of satisfaction and acceptability.

4.4.1 *Coefficient of subjects' satisfaction overall*

Subjects in both groups were satisfied with their performance, some much more so than others. The mean coefficient of subjects' satisfaction in direct translation was as follows: translators (.84); teachers (.76). In inverse translation, the mean coefficient was as follows: translators (.73); teachers (.76).

Taking all subjects with a coefficient of satisfaction of 1, it may be observed that a total of twenty subjects had a coefficient of 1 in direct translation (15 translators and 5 teachers). Of a total of 12 subjects with a coefficient of 1 in inverse translation, 6 were translators and 6 were teachers. All translators with a coefficient of 1 in inverse translation had a coefficient of 1 in direct translation, whilst only one teacher (PF7) had. It would thus appear that translators have more clearly-defined, homogeneous criteria that inform their perception of the acceptability of their translations.

4.4.2 *Coefficient of satisfaction for each Rich Point*

Subjects' coefficient of satisfaction is similar for each Rich Point (cf. Table 9). Satisfaction would therefore not appear to depend on the nature of the problem solved.

There would appear to be only a slight difference in inverse translation (RP4: *geografía comarcal*, problem of intentionality). Although a similar number of subjects in both groups identified this Rich Point as a problem (translators: 68.6%; teachers: 62.5%), teachers were more satisfied (.85) with their solutions to the problem than translators (.61).

Table 9. Coefficient of satisfaction for each Rich Point

	Direct translation		Inverse translation	
	Translators	Teachers	Translators	Teachers
RP1	.78	.76	.78	.79
RP2	.76	.61	.82	.74
RP3	.89	.90	.74	.70
RP4	.83	.64	.61	.85
RP5	.89	.76	.69	.76

4.4.3 *Coefficient of satisfaction and acceptability*

Finally, the coefficient of subjects' satisfaction was compared with the index of acceptability to determine whether or not subjects' satisfaction had a bearing on the acceptability of their solutions to translation problems.

There is a marked dispersion of subjects in both inverse and direct translation (cf. Figures 2 and 3). Some subjects cannot be seen because they overlap.

No relation between subjects' satisfaction and the acceptability of their solutions is observed in either of the two groups.

To confirm this finding, the coefficient of satisfaction was compared with the index of acceptability in those subjects that had obtained the highest (top 25%) and lowest (bottom 25%) indices of acceptability. A high index of acceptability for translators was set at .90–1.00 (direct translation) and .80–1.00 (inverse translation). A low index of acceptability for translators was set at .60–.30 (direct translation) and .30–.10 (inverse translation). In the case of the teachers, a high index of acceptability was set at .70–.90 (direct and inverse translation); a low index of acceptability was set at .30–.00 (direct and inverse translation). No significant difference was found in direct or inverse translation between translators or teachers with high or low indices of acceptability. This would confirm the result already obtained, i.e. that no relation exists between subjects' satisfaction and acceptability.

Moreover, a comparison between the acceptability of the solutions of translators with a coefficient of satisfaction 1 (very satisfied with their solutions) and the acceptability of the solutions of all translators in both inverse and direct translation confirms

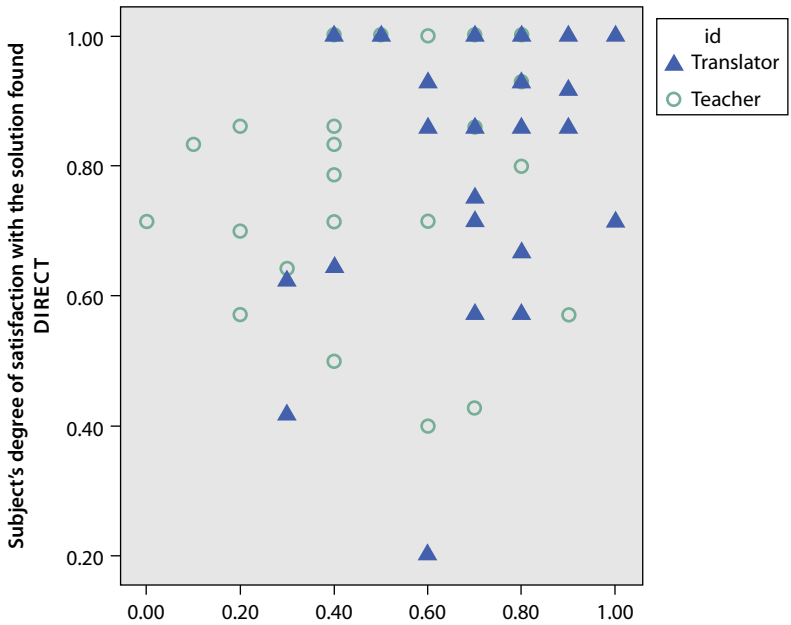


Figure 2. Coefficient of satisfaction and acceptability (direct translation)

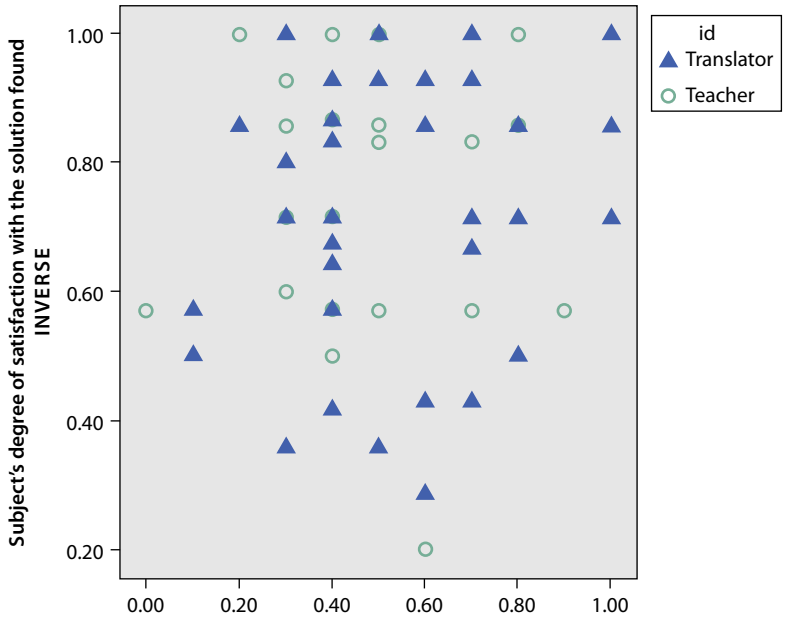


Figure 3. Coefficient of satisfaction and acceptability (inverse translation)

that there is no correlation between subjects' perception of the acceptability of their solutions and true acceptability (cf. Table 10).

A high degree of satisfaction does not guarantee the acceptability of solutions.

Table 10. Comparison of satisfaction and acceptability in translators

Acceptability	Direct translation	Inverse translation
Translators (coefficient of satisfaction 1)	.76	.52
All translators	.73	.52

5. Conclusions

In the first place, results of particular significance for our research design were obtained:

1. The homogeneous results obtained for the coefficient of perception of the overall difficulty of the text in direct and inverse translation leads us to conclude:
   a. The homogeneity of the texts used for direct translation in our experiment: the difficulties presented in the texts appeared to be similar. The way in which comparable texts were selected in the three languages for direct translation was thus deemed appropriate, since translators perceived the texts to be of comparable difficulty.

b. The homogeneity of the sample of translators: there was no language-dependent difference in the perceived difficulty of the texts to be translated either in direct or inverse translation. The questionnaire used to select subjects for the experimental group of translators was therefore appropriate.

2. The Rich Points selected were appropriate since all subjects found them difficult to translate, both in inverse and direct translation, and described them in terms of different categories (multidimensionality of translation problems).

Regarding translators' and teachers' behaviour in relation to translation problems and their ability to identify and solve them, we have found, first, that each group differs in their perception of the overall difficulty of the texts. In direct translation, the group of teachers perceived the text to be more difficult than the translators, possibly due to their inexperience in text production in their mother tongue, and their inexperience in translation. In inverse translation, the teachers (with the exception of the teachers of French) also perceived the text to be more difficult than the translators. In both groups the perceived degree of difficulty of the text was much higher in inverse translation. This would indicate that directionality plays a role in the difficulty of translation problems.

No relation, however, was found between subjects' perception of the overall difficulty of the text and the acceptability of their solution to translation problems. Nor was a relation found between subjects' satisfaction with their solutions to translation problems, i.e. personal perception of the quality of the translation, and real acceptability. We may therefore conclude that subjects' perceptions of the difficulty of the translation of a text and their perceptions of the quality of their performance, have to do with personality traits such as self-esteem, self-criticism, and so forth – psycho-physiological components in PACTE's theoretical model (PACTE 2003).

Moreover, the problems identified both by translators and teachers varied greatly depending on the individual. This would appear to point to subjectivity when facing prototypical problems. This may well have to do with the distinction between difficulties in translation (subjective) and translation problems (objective) proposed by Nord (1991).

Finally, on the basis of the results obtained, we cannot affirm that the characterisation of translation problems is a feature of translation competence. Rather, the ability to explain the nature of a problem would pertain to the field of *explicative knowledge*, characteristic of the theorist in Translation Studies. It has to do with the distinction made in cognitive psychology between explicative knowledge (*know why*) (cf. Wellington 1989; Pozo and Postigo 1993), which identifies theoretical knowledge; declarative knowledge (*know what*) and operative knowledge (*know how*) (cf. Anderson 1983); and conditional knowledge (*know when and why to use knowledge*) (cf. Paris, Lipson and Wixson 1983; Paris, Cross and Lipson 1984; Schraw and Moshman 1995). A characteristic of translation competence would therefore be to possess declarative knowledge specific to translation (cf. PACTE 2008), as well as operative and conditional knowledge, i.e. *know how* to identify and solve translation problems by applying the relevant knowledge and strategies.

Nevertheless, the results obtained have yet to be triangulated with results of other variables, in particular, in relation to the translators who had obtained the highest score in different aspects of our study (acceptability of solutions, knowledge of translation, decision making, use of instrumental resources, and so forth.).

References

Alves, F. 1995. *Zwischen Schweigen und Sprechen: Wie bildet sich eine transkulturelle Brücke? Eine Analyse von Übersetzungsvorgängen zwischen portugiesischen und brasilianischen Übersetzern.* Hamburg: Kovac.

———. 1997. "A formação de tradutores a partir de uma abordagem cognitiva: reflexões de um projeto de ensin." *TradTerm. Revista do Centro Interdepartamental de Tradução e Terminologia* 4 (2): 19–40.

Anderson, J. R. 1983. *The Architecture of Cognition.* Cambridge, Mass.: Harvard University Press.

Bachman, L. F. 1990. *Fundamental Considerations in Language Testing.* Oxford: Oxford University Press.

Bell, R. T. 1998. "Psycholinguistic/cognitive approaches." In *Routledge Encyclopedia of Translation Studies*, M. Baker (ed.), 185–190. London: Routledge.

Catford, J. C. 1965. *A Linguistic Theory of Translation: An Essay in Applied Linguistics.* London: Oxford University Press.

Giegler, H. 1994. "Test und Testtheorie". In *Wörterbuch der Psychologie*, R. Asanger and G. Wenninger (eds), 782–789. Weinheim: Psychologie Verlag Union.

Hönig, H. G. and Kussmaul, P. 1982. *Strategie der Übersetzung.* Tübingen: Gunter Narr.

Hurtado Albir, A. 2001. *Traducción y Traductología. Introducción a la Traductología.* Madrid: Cátedra.

Jakobsen, A. L. 2002. "Orientation, segmentation, and revision in translation." In *Empirical Translation Studies: Process and Product [Copenhagen Studies in Language Series 2]*, G. Hansen (ed.), 191–204. Copenhagen: Samfundslitteratur.

Kiraly, D. 1995. *Pathways to Translation. Pedagogy and Process.* Kent: The Kent State University Press.

Krings, H. P. 1986. *Was in den Köpfen von Übersetzern vorgeht. Eine empirische Untersuchung zur Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern.* Tübingen: Gunter Narr.

Lörscher, W. 1991. *Translation Performance, Translation Process, and Translation Strategies. A Psycholinguistic Investigation.* Tübingen: Gunter Narr.

Mounin, G. 1963. *Les problèmes théoriques de la traduction.* Paris: Gallimard.

Nord, C. 1991. *Text Analysis in Translation.* Amsterdam: Rodopi.

PACTE. 2002. "Exploratory tests in a study of translation competence." *Conference Interpretation and Translation* 4 (2): 41–69.

———. 2003. "Building a Translation Competence Model." In *Triangulating Translation: Perspectives in Process Oriented Research*, F. Alves (ed.), 43–66. Amsterdam/Philadelphia: John Benjamins.

———. 2005a. "Primeros resultados de un experimento sobre la Competencia Traductora." In *Actas del II Congreso Internacional de la AIETI. Información y documentación*, 573–587. Madrid: Publicaciones de la Universidad Pontificia Comillas, CD-ROM.

———. 2005b. "Investigating translation competence: Conceptual and methodological issues." *Meta* 50 (2): 609–619.

———. 2007a. "Zum Wesen der Übersetzungskompetenz. Grundlagen für die experimentelle Validierung eines Ük-Modells." In *Quo vadis Translatologie? Ein halbes Jahrhundert universitärer Ausbildung von Dolmetschern und Übersetzern in Leipzig. Rückschau, Zwischenbilanz und Perspektiven aus der Außensicht*, G. Wotjak (ed.), 327–432. Berlin: Frank & Timme.

———. 2007b. "Une recherche empirique expérimentale sur la compétence en traduction." In *Actes du Colloque International: Quelle formation pour le traducteur?,* D. Gouadec (ed.), 95–116. Paris: La Maison du Dictionnaire.

———. 2008. "First results of a translation competence experiment: Knowledge of translation and efficacy of the translation process." In *Translator and Interpreter Training. Issues, Methods and Debates*, J. Kearns (ed.), 104–126. London: Continuum.

———. 2009. "Results of the validation of the PACTE translation competence model: Acceptability and decision-making." *Across Language and Cultures* 10 (2): 207–230.

———. 2011. "Results of the validation of the PACTE translation competence model: Translation project and dynamic translation index." In *IATIS Yearbook 2010,* S. O'Brien (ed.), 30–53. London: Continuum.

Paris, S. G., Lipson, M. Y. and Wixson, K. K. 1983. "Becoming a strategic reader." *Contemporary Educational Psychology* 8: 293–316.

Paris, S. G., Cross, D. R. and Lipson, M. Y. 1984. "Informed strategies for learning: A program to improve children's reading awareness and comprehension." *Journal of Educational Psychology* 76 (6): 1239–1252.

Pozo, J. I., Gonzalo, I. and Postigo, Y. 1993. "Las estrategias de aprendizaje como contenido del currículo." In *Estrategias de aprendizaje*, C. Monereo Font (ed.), 47–64. Barcelona: Domènech.

Presas, M. 1996. "Problemes de traducció i competència traductora. Bases per a una pedagogia de la traducción". (Doctoral dissertation, Universitat Autònoma de Barcelona, Spain.)

Scavee, P. and Intravaia, P. 1979. *Traité de stylistique comparée du français et de l'italien.* Paris: Didier.

Schraw, G. and Moshman, D. 1995. "Metacognitive theories." *Educational Psychology Review* 7: 351–373.

Vázquez Ayora, G. 1977. *Introducción a la Traductología.* Washington: Georgetown University Press.

Vinay, J. P. and Darbelnet, J. 1958. *Stylistique comparée du français et de l'anglais. Méthode de traduction.* Paris: Didier. Translated into English by J. Sager & M. J. Hamel, 1995: *Comparative Stylistics of French and English. A Methodology for Translation.* Amsterdam/Philadelphia: John Benjamins.

Wellington, J. J. 1989. *Skills and Processes in Science Education. A Critical Analysis.* London: Routledge.

Wilss, W. 1996. *Knowledge and Skills in Translator Behaviour.* Amsterdam/Philadelphia: John Benjamins.

Appendix 1: Texts

Direct Translation Text

1Email virus strikes in new form

Computer users were warned last night to be on the lookout for an email virus that can steal confidential information and allow hackers to take control of infected machines. The virus, a new variant of the BugBear email worm that infected tens of thousands of computers around the world last October, began to spread rapidly from Australia to Europe and the USA at around 8am yesterday. According to MessageLabs, a **5Cheltenham-based virus filtering firm** which reported about 30,000 infected messages in 115 countries, the **3propagation rate** of BugBear. B almost **3doubled** every hour throughout the morning. There was also a huge **3surge** as US users came online. Like its predecessor, the variant spreads by sending itself as an attachment to every address in an infected machine's email address book. To disguise where it came from, it uses different subject headings. As well as searching for anti-virus software and disabling it, BugBear.B installs a **2keylogger** to record what the user types, which may allow hackers to record confidential information such as credit card details and passwords. It also installs a **4 "Trojan horse" program which could allow a hacker to take remote control of infected machines**. […]

The Guardian – Friday, June 6, 2003

Inverse Translation Text

La Plana Novella

La Plana Novella es una antigua heredad adquirida por el **1Indiano** Pere Domenech i Grau en 1885 que se encuentra en una pequeña planicie en el centro del Parc Natural del Garraf y pertenece al municipio de Olivella. La Finca fue declarada colonia agrícola 10 años más tarde por el **2gobierno alfonsino**, pero de aquella época perdura una leyenda de **3desenfreno y dilapidación** que hizo desaparecer la **1fortuna del americano** . El estilo arquitectónico del Palacete es ecléctico, es decir que mezcla diferentes estilos. **4La geografía comarcal de Cataluña** lo califica de "Castillo de Bambalinas" como si fuese un decorado de teatro. Sin ningún tipo de duda la construcción estilísticamente más original de Palau Novella es el lavadero gaudiniano, pero una de las piezas más características y llamativas del Palau es el **5común**, conocido como **5"la trona"**.

http://www.laplananovella

Appendix 2: Questionnaire on translation problems

1. How difficult do you think this text is to translate?
   On the scale between "very easy" and "very difficult", put a cross on the line of squares below to show how difficult you think it would be to translate this text.

   Translation of this text is very easy — Translation of this text is very difficult

   ❑

2. What are the general characteristics of the text that make you think so?

3. What were your priorities when translating the text?

4. What were the main problems you found when translating this text? Name 5 and answer the following questions about each.

Problem 1: ……	Why was it a problem?
	What were your priorities when solving it?
	Explain as clearly as possible what you did to solve it.
	Are you satisfied with the solution? ❑ Yes ❑ No Why?

"This led me to start thinking about how this happened, and what the process behind it would be"

An interview with Professor Birgitta Englund Dimitrova*

Elisabet Tiselius
University of Bergen, Norway / Stockholm University, Sweden

On becoming a translation scholar

Elisabet Tiselius: *Birgitta, you're the first professor of Translation Studies, in Sweden, aren't you?*

Birgitta Englund Dimitrova: Yes, although formally I'm appointed "professor with a focus on interpreting and translation". It's called that because the position I had initially as a lecturer was formally called "lecturer with a focus on interpreting and translation". So, the professorship isn't formally in Translation Studies. It would have been better if it had been called that, but that's the way it is.

---

* In preparing this tribute to Professor Englund Dimitrova, it was our intention as editors of the volume not only to surprise her with the publication, but also to hear Professor Englund Dimitrova's own thoughts on some of the topics discussed herein. We hope that this interview will accomplish this. As for the surprise factor, the interview was contextualized as a publication for a Swedish popular science journal (*Språk*) focusing on linguistic and language issues, and Professor Englund Dimitrova kindly agreed to participate in the project. This text is the outcome of two interviews of Professor Englund Dimitrova by Elisabet Tiselius that took place in the late summer and early autumn of 2010, one in Stockholm and one in Leuven. The discussion was of a fairly general nature and at one point, as Elisabet tried to steer into more specific issues related to process research and to research methodology, Birgitta suddenly asked: "But do you really think the readers of this journal are that interested in methodology?". We hope that we have succeeded in giving Birgitta a voice and believe that in reading her life story, one is reminded of the challenges to Translation Studies research as a field and to Translation Studies scholars individually. The following excerpts from the interview were translated from Swedish.

E: *Is there a difference?*

B: No, there is no difference; it has more to do with the use of the term Translation Studies. For me it has become a strategic decision. For a long time I used "research in interpreting and translation" instead. I used this sort of more descriptive terminology and I did in some way avoid the term *översättningsvetenskap* (Translation Studies), but some years ago I simply decided that from now on I'll call our field Translation Studies and if I need an explanation I'll add "research in translation and interpreting". I think you need a short and distinct term for the area you would like to make into a discipline, and then *översättningsvetenskap* is, I think, the best term in Swedish for what in English is called *Translation Studies*, and *traductologie* in French, and *Übersetzungswissenschaft* in German.

E: *When did you become a professor?*

B: I became a professor in 2006.

E: *And why is it that you have become such an advocate for Translation Studies, and a pioneer in that field?*

B: Well, it has been a long road. I didn't begin my academic career with translation or Translation Studies, let alone with interpreting. I am a Slavist and earned a PhD with a thesis on Bulgarian and Macedonian syntax, something like contrastive linguistics you could say. And in those days – it sounds like it was a really long time ago, but this was in the seventies – and in those days research on various aspects of translation was not really accepted in the language departments. There were of course a few pioneers, but I was by no means one of them. I was occupied with other things. But then thanks to my knowledge of Bulgarian, which is a rare language in Sweden, I was hired for various assignments as a translator and interpreter. Not that I tried to get such assignments, it was rather that people needed someone who spoke Bulgarian. This meant that I gained some experience in interpreting that way, but really on a very practical level. At that time I had a position in the department of Slavic languages at Stockholm University and teaching Bulgarian and doing research. But that was a short term contract, as many academic positions tend to be, unfortunately. And when the contract ended I was forced to leave the academic world and the research I was doing then.

By chance, I got a position at the Stockholm County Council in the Public Health Care Services Committee, where they had a project in the eighties reviewing the interpreting services for immigrants in the health care sector. They were looking for a number of immigration advisory officers who would work with this project. I applied and got the position in the Southern Health Care Provision of the Stockholm County Council. I probably got it mostly thanks to my (limited) experience of interpreting and translation. I managed to convince them that the fact that I held a PhD was not a disadvantage. I remember mentioning that in my written application. I tried to explain that writing a thesis could be something positive outside the academic community too. Obviously I managed to convince them.

The position didn't involve any research at all. It was a very hands-on review procedure for interpreting services in the health care sector. I would make suggestions on how these procedures could be improved, such as improving the use of interpreters by training the staff, the users, and providing them with information and so on. It was a very exciting job, and although I had some experience as an interpreter, I learned a lot, both about interpreting and about the work in the health care sector. It was fascinating. We were several immigration advisory officers in the different Health Care Provisions in the Stockholm County Council. We organized for example staff training, and that was very interesting. Back then in the eighties there were no researchers in interpreting, but researchers in other fields lectured for us. They talked about non-verbal communication and cultural encounters and as they were good researchers they brought hand-outs and reference lists. My dream of working with research which had been there ever since high school, and which I had been lucky to pursue for a couple of years in Slavic languages, was apparently still there. I remember that the first summer after I had started at the County Council, this was before I had children of course, I spent part of my holiday at the National Library reading up on non-verbal communication and cultural encounters in interaction, the reference lists from the staff training were an initial help in my search for literature. And then I began to think, well, this is interesting for interpreting. I read and took notes, just because I found it so fascinating. I didn't know then whether it could lead to something more, so to speak. I just did it. It sort of came from within.

E: *Is this when the pioneer was born?*

B: Yes, maybe you can say that. Then a couple of years went by, and the project with the County Council came to an end. Then one day in the spring of 1986, when I had had my first child, I received a letter. It was from Stockholm University and from the Department of Slavic languages, I could see that on the envelope, but there was no note inside; the only thing that was included in the letter was a copy of an announcement for a post-doc position in bilingualism with a focus on interpreting and translation – talk about a long name! The deadline for application was three weeks later. To this day, I don't know who sent it to me. The requirement was a PhD in modern languages and I fulfilled that requirement, then there was really nothing else that was required. I don't recall if you had to have experience in translation or interpreting, or have had done research in those fields. In those days I don't think it was a relevant requirement, since very few people would have been able to apply in that case. I had little Nikolaj, who was a few months old, and I had actually already started working too, I was a substitute lecturer in Bulgarian at Uppsala University. I still managed to produce an application. I found the notes I had made a few years earlier at the National Library and I thought I had to present a project, it was not enough just to apply for the job without a project. So, I made a project plan based on interaction and non-verbal communication in interpreted encounters and applied with that. I don't know if that was the trick that did it or if it was the overall application, but I got the job anyway. And I started on December 1st, 1986.

E: *But there was no predecessor? Had somebody had a position focusing on interpreting and translation before?*

B: Well, of course there were researchers at Stockholm University who had a great interest in translation. For instance, in the German department there were many people who had had a significant interest for many years in translation, mainly literary translation, but also, specialized language translation. Also, in the Department of Romance Languages and in the Department of Slavic Languages, as well as other departments some researchers had worked with translation, and there had been an interest in translation issues. But I don't think anybody had shown an interest for more modern areas of Translation Studies or interpreting.

E: *No, and I suppose that it was also an important factor that there was actually a label put on it? It was interpreting and translation independent from any particular language or other discipline.*

B: Yes, of course. This may of course also have had something to do with the fact that the Institute for Interpretation and Translation was established in 1986. The Institute itself was not planned to have any research, and the position I got was a post-doc position. But I believe Stockholm University wanted to profile itself by investing in interpreting and translation research and thereby creating a position earmarked for that. The position would not be located at the Institute, but instead the holder would belong to the department where she had her major subject. When I started on December 1, 1986, the Center for Research on Bilingualism was not yet established so I was placed in the Slavic department. I was then pregnant again and I was on sick leave most of the pregnancy, and then I was on maternity leave, so I didn't take up the position until the spring of 1988.

By then, the Center for Research on Bilingualism had been established and Kenneth Hyltenstam, the director, had asked me then what I wanted to do, since my position was called post-doc in research on bilingualism focusing etc. etc., so he had asked me if I wanted to remain in the Slavic department or if I wanted to come to the Center. I thought it would be interesting to get some new impulses so I answered that I would like to come to the Center, and that was it.

E: *This is when you started to do research on interaction?*

B: Yes, exactly, I went out into the bush, not literally of course, but in any case to a Swedish health care center, and I brought a video camera and recorded two interpreter-mediated medical encounters that I analyzed.

E: *At this time there were no other researchers around you who were also working with Translation Studies?*

B: No, none.

E: *And if you look at researchers at the international or Nordic level?*

B: There were not many around, I had met Cecilia Wadensjö, who began her PhD studies then. I think I met her in the spring of 1987, when she was in the beginning of her research studies at Linköping University, and I realized that we had very similar research interests such as interaction and so forth. But otherwise I didn't know of anybody.

One thing that I regret, or maybe regret is a strong word, but one thing that I didn't understand then was how important it was to reach out. If I was one of the first ones, I had to reach out. That's what it was all about, really, reaching the international community with what I did. The study I did on the interaction in interpreter-mediated medical encounters, for instance, was published in Swedish as a report in the Centre for Research on Bilingualism series – "Reports on Bilingualism". I can think now that it was a bit unwise because I had actually been the first to do things that others have since then done. I mean we have not used the same theoretical grounding, so I don't necessarily mean that I would have made a breakthrough, or how should I put it… But it highlights how difficult it is, and how important it is for us as researchers to really consider what the best channel for publishing is.

E: *You have become quite influential in process research. I believe that one of the major reasons why you have become so influential is your major study on process that was reported in your book "Expertise and Explicitation in the Translation Process", it laid the foundations for this. But when did you come in contact with international researchers? Because you have told me about the work at TÖI (the Institute for Interpretation and Translation Studies) and research on interpreting and so forth, and this has been very Swedish-based, but then there's some springboard somewhere.*

B: Yes, that has come gradually, partly through conferences where I had the opportunity to participate. The first translation conference that I remember was the Scandinavian Symposium on Translation Theory, which was held a couple times in the late eighties and early nineties. I think it was in Oslo in 1988 or 1989 when I first had the opportunity to participate, and there were a number of "names" present which I then had the opportunity to follow, so to speak. That's the way it is when you become interested in an area and write about it and lecture about it, and search for literature about it. After a while you get into it, first the literature and eventually you meet and get to know people. I cannot really say that I went to a particular place and suddenly realized there were others. But I have, however, experienced this another time, in the beginning of my doctoral studies. I thought that I was the only one who was interested in Bulgarian. I thought it must be very unusual until I went to a summer course in Macedonia. It was the summer of 1969 and I realized that there were several others who were interested in Bulgarian, including a guy from Norway that I'm still in touch with. In that case it was a revelation, but for Translation Studies it was not in the same way. It was a gradual process.

E: *You described how you started with research on dialogue interpreting in a very concrete way. But you have also spent a lot of time and effort on translation.*

B: Yes, certainly, and that started off in exactly the same way where I had a few thoughts on doing some form of statistical analysis of translations in different directions, counting the number of words and so on. And it led to an article, or a study, that led me further because it made me aware of common features in translation between different languages, but also differences between different translators in fact. This is something that still interests me very much; to what extent the features we see in a translation are more general or universal or to what extent it is even on an individual level. Individual differences between translators, and to what extent these differences and similarities depend on, for example, how much experience you have of translation. I was also interested in what and how much you can actually perceive of the process when you start studying it.

I started just as so many others do: I studied a source text and compared it to a target text. And then you see that some things happened here. This led me to start thinking about how this happened, and what the process behind it would be. And that has become my main area of interest when it comes to translation.

On process research and methodology

E: *Did you start to think about this on your own? Or was it other researchers or other things you read that got you interested in the process?*

B: No, on the contrary, these were very much my own thoughts. I have to say that I get very annoyed when I see studies where people have just started from a source text and a target text, and from this argue that the translator has thought this or that. It has always upset me since I believe that you cannot get absolute proof of the translator's actual thoughts from the material. You can of course assume certain things, but the material itself says nothing about what the translator actually meant. This is what I have mainly tried to examine in my principal study on process, but it's still very difficult. It is difficult and you will never get complete access to thoughts and ideas, but you can approach them a little bit more anyway.

E: *Why is it so important to study the process?*

B: In my opinion it can lead to a deeper understanding of what translation and interpreting looks like. I also think that from an educational perspective it's absolutely essential to know as much as we can about the process, both in order to plan the teaching in a good way, but also to focus on the process and not just fill the classes with texts that students have to translate and tell them that here you translated like this and here you translated like that, but instead actually help students to gain insight into their own process, because processes may vary between different people, and you have to raise that awareness with the student. I think that is very important not only to students but also for professional translators and interpreters.

E: *And how would you define an expert, a concept that also has been important in your work?*

B: The definition of an expert depends on the purpose of calling someone an expert, if you work within a certain paradigm, the expertise approach, for instance, then you have one idea of how to define an expert. The purpose may be very different if you use "expert" as some sort of general concept. I have adopted the expertise paradigm in such a way that in my principal study I didn't define an expert, I never concluded that any of my subjects were experts. I labeled them senior and junior professionals, where the seniors had more than 15 years of experience and then I discussed my results in terms of what might be expert characteristics in their performance. But a definition of an expert also implies quality, and since there is no unity in the definition of quality in Translation Studies, how can we then identity for instance the consistently outstanding performance, which is a central concept in the expertise approach. If we cannot even define what an outstanding performance is, then how could we say that it's consistent? This is also evident when you look at revisions of translations, which also could be one way to identify quality in translation. In studies done on revision, the subjects do not agree on what an outstanding text should look like. So not even on the shop floor is there a common view on quality. Therefore, it is of course easier to define expertise in terms of short distance runners or Wimbledon winners, where the outstanding performance is less disputable.

E: *You said that you're very interested in the process and that you become very annoyed if conclusions are drawn about processes without really knowing anything about them. How do you see this process? What is your opinion about it? I mean, it's easy to say that you shouldn't draw conclusions about the process without knowing anything about it, but what can you know about it and how can you know something about it?*

B: Oh, what a difficult question! Well there's quite a lot you can know and, above all with the combination of different research methods, you get a better basis for drawing conclusions than the basis you have if you only use source text and target text. One of the aspects where I have had a particular interest is in revision as part of the process. These revisions can be studied through key-logging, and also with the help of think-aloud protocols and perhaps also with retrospection. This gives us some sort of general, I won't say universal, but in any case a very common pattern. And this pattern shows that translators start with something that is very close to the source text and then you gradually move away from it in order to come closer to something more acceptable from a target language perspective. But this is an ongoing process in the brain and the possibilities to study the brain are not so easy accessible.

E: *Is process research a strong field within Translation Studies?*

B: Yes, stronger than I think I understood earlier. There's a lot of good, well-supported research going on, and I'm happy to see how many we are. If I compare to when I started in the early 1990s, when I started thinking that you cannot draw conclusions

on the process by only studying the product, and I started exploring the process using key-logging with Scriptlog a software from Gothenburg University (they had developed software, not for translation, of course, but still), and I also started exploring think-aloud protocols. And I had the idea that this could enable me to follow the process in real time, I would be able to see how the different translation solutions were developing. This was extremely exciting, but also very lonely, there seemed to be no one else, or at least very few, even internationally. Only gradually did I realize that there actually were other people working with the same issues, at the Copenhagen Business School for instance, in Finland, in Germany but also others.

E: *And today you see the result of these early efforts?*

B: Yes, and it's fantastic that all these projects today are combining methods in similar ways that I started exploring back then with Scriptlogs, think-aloud protocols and interviews, and other methods too of course.

E: *And what would you say are central themes today in process research? Which trends can you see there?*

B: Many researchers are interested in longitudinal studies on how translation competence develops over time and what can be said to characterize translation competence. This was also one of my early areas of interest: the development of translation competence. And also the things that are studied at Copenhagen Business School right now, they are exploring very basic aspects of the translation process, and apart from the keylogging software Translog, they also use an eye-tracking device which gives very interesting results on for instance how you read for translation as opposed to how you normally read. These are new methods that provide new possibilities to analyze the translation process, which then leads back to basic questions on the translation process. How does the process actually happen?

E: *Do you mean that with by new technology we can question old results?*

B: Either that, or the fact that new technology allows us to dig deeper into earlier results.

E: *Are there any research methods that are particularly relevant or good to process research?*

B: It's very difficult to answer that question. It depends on the research question, the researcher's personality and so forth. If we take think-aloud protocols for instance, they are very popular, but they have an innate vagueness that needs a lot of interpretation, and that's not suitable for everyone. Some methods are good because they are easy to use, or give immediate results. But I don't have any pet methods. Each method has its advantages and disadvantages. Take a corpus for instance, a corpus is a wonderful tool if you search for one very specific phenomenon. But if you're interested in metaphors,

for instance, which are less specific, it may be very difficult to use a corpus to find metaphors.

E: *Process research in Translation Studies is sometimes criticized from a more psychologically oriented perspective for having too small samples or not being strict enough about controlling variables and so forth. On the other hand it is also criticized from a more sociologically oriented perspective for not being ecologically valid and not taking the real event into consideration. We are somewhere in between, not really one, but not really the other either. Is there a way around that? Is there a disadvantage being in between? Can we accept that?*

B: Maybe we have to accept it. If we were to apply the methods we borrow from other disciplines much more rigidly, then we may have to accept that we get further away from each other. Process researchers would be confined to the white room with very large samples and the question is whether we could even find enough subjects for such experiments. We try to balance this as process researchers in Translation and Interpreting Studies. We try to design our experiments to be as rigid as possible in terms of controlling variables, but also to have some ecological validity, so that the experiments will tell us something about the reality we're trying to interpret. And maybe this is an innate contradiction that we have to live with if we want to be part of the discipline of Translation Studies.

On the discipline of Translation Studies and the TS research community

E: *I'll change topic here. Today there is one more Professor in Translation Studies in Sweden. And as I've said, I see you as a pioneer, at least in Sweden. So I would like to ask you very broadly: Has this been a very deliberate effort? You said that when you started using the term översättningsvetenskap (Translation Studies), it was a conscious decision. But have you also deliberately worked to establish Translation Studies as a field in Sweden, or was it a natural process through the creation of training programs, seminars and so forth?*

B: No, I cannot say that I had a vision from the beginning that I would create Translation Studies as a field in Sweden – I think it would be very presumptuous to say that I had such a vision from the beginning. It's a vision that has developed gradually and in particular after I began at TÖI (the Institute for Interpretation and Translation Studies), when I was asked, already in 1994 or perhaps in early 1995 to develop a research and development program for TÖI. At that point I had to start thinking ahead. I cannot remember if I wrote something about Translation Studies as a discipline at that time, but the specific goal to get a PhD program in the subject was there already then. This was adopted by the Board of TÖI in December 1995, 15 years ago now, so it's high time to implement it.

But no, when I started my post-doc position, I had my roots in Slavic linguistics after all. Especially since I hadn't worked at Stockholm University for a few years, it's clear that you cannot suddenly say that now I'll create a new discipline. However, I do admit it: my goal back then was to try to ensure that I obtained a permanent position, since the post-doc position was not a permanent one. I thought to myself that I would not go through that once again, since I had already once before been in a situation where a short-term position came to an end and I had had to leave. So I was a bit strategic, and thought that if there's going to be another position created for this area, besides the post-doc position, it seems likely that it will be focused on both interpreting and translation, it will not be one or the other. So for this reason I thought it would be best to publish a bit on both. That way I can say that my goals in the beginning were sadly selfish.

E: *They still led you onto something, didn't they?*

B: Yes.

E: *And do you think that there is a lively and good research community of Translation Studies in Sweden?*

B: Yes, there is. Many worthwhile things are accomplished, but most research projects are, unfortunately – don't get me wrong – but most research projects are still made by doctoral students. Of course, there's nothing wrong with that, when I say unfortunately I don't mean that such research is less valuable. In many disciplines the doctoral students' research is an important part of the research discipline. But in this case – since we are so fragmented and graduate students belong to different departments and different subjects, often without contact with each other – it means that they are all in danger of starting out by re-inventing the wheel. Whereas if you have a more well-defined research community, graduate students would be able to start further down the road and thereby reach further with their theses. But plenty of interesting work is being done. If the research is vibrant – well, it's a financial issue, but it's obviously the same in all disciplines. We need a closer contact between those who are doing Translation Studies, we definitely need that.

E: *In Sweden or in the Nordic countries?*

B: Yes, in Sweden – the Nordic countries too, of course, but in Sweden, because the question concerned Sweden. We try, of course. We do have our research seminars here at TÖI, and researchers from different departments, mostly from Stockholm perhaps, come and present their work. But I feel that it's still a bit divided.

E: *If we take a look at the international level, do you think that there's a lively research community in the areas that you're interested in, such as process research or dialogue interpreting?*

B: Yes, I think there is, actually, within both areas. There may not be a very large community of researchers in process research, but we know each other, it's a lively community and, yes, I certainly think that there's a sense of intelectual fellowship. Within dialogue interpreting, the field has indeed benefited from the Critical Link Conferences, which have been a joint effort from the profession, authorities and researchers in this area. Research in dialogue interpreting is younger than most areas in Translation Studies, but has developed into a growing and vibrant research community.

E: *It is a field that has grown?*

B: Yes, absolutely.

E: *You came to TÖI as a lecturer in 1994, and then you became a professor in 2006.*

B: Yes, I was appointed professor shortly after I finished my principal study on explicitation and published it with Benjamins.

E: *And what have your tasks been at TÖI? You have taught translation, of course.*

B: Yes, I have taught translation theory and I have developed many new courses. I have actually been involved to a greater or lesser extent in all the courses we give now, but mostly for translation.

E: *We discussed earlier the fact that Translation Studies has become a discipline during your time as an active researcher. But what does it mean to the discipline in Sweden, for instance, that there are no PhD students within that discipline?*

B: That's what you need to ask yourself. It leads you on to issues such as, What is a discipline? When can you call a field of research a discipline? And the fact that you have a PhD program is one of the factors creating a discipline. Traditional disciplines have PhD programs, and it can be difficult to introduce a new PhD program in a new discipline. The university decides whether a new PhD program should be established, we cannot decide that for ourselves, and the university has to have a general overview.

In some respects you can say that Translation Studies has become a discipline. Again, it is a question of how we define a discipline. What is required before we can say that this is a discipline? And in some respects a PhD program may be one of the conditions, or perhaps such a simple thing as registering your publications in a university database. Many university databases do not have Translation Studies as a subject. So when you register your publication, you have to register it as belonging to some other discipline, linguistics or bilingual studies or something else. And that way you're made invisible. But a positive development in Sweden is that more and more students are writing their BA theses in Translation Studies, and also focusing on interpreting, which means that they get their BA in Translation Studies, and the same thing goes for the Masters dergree. I think this is a very positive development, in particular the fact that we have students writing on interpreting. This way we build Translation Studies from the ground.

E: *We also talked about the international level earlier, and in Translation Studies the picture is often the same internationally. Translations scholars often belong to other departments than Translation Studies since such departments don't exist. Does the problem lie there, in the fact that there are no undergraduate courses?*

B: Well, there are undergraduate courses at several schools. This is what has happened in recent years, undergraduate courses have grown very strongly, and also at a more advanced level there are many courses in both Europe and elsewhere. But at the next level, yes, the issue of PhD programs is an important one; it concerns both the status of the discipline and its visibility. But this is something we share with other new subjects. The difference for us is that we lack support from the authorities, which many other new subjects tend to receive. In Sweden, for instance, there's the case of Caring Sciences, which only thirty years ago wasn't a research subject, and the fact that it has grown into a research discipline is because the Swedish authorities believed strongly that professional training and education in different areas should be linked to research. But for some reason this isn't true for Translation Studies.

E: *What areas are important to conduct research on in Translation Studies? What issues should we pursue?*

B: I still believe very stubbornly that many of the answers lie in Process Research, and PR in a very wide respect, not only looking at the process of translating and interpreting, but also the process when a reader reads the translation. How do you read a translation? And what does the process look like when you listen to an interpreting? And also didactic questions: What is the best possible way to learn translation and interpreting? How do we teach it in the best possible manner? Are the methods we use the best ones? And, of course, all the issues in dialogue interpreting, which is still a field with many potential research opportunities.

E: *I would like to ask you about trends in Translation Studies. There are ongoing discussions in Translation Studies about turns and paradigms. Would you say that Translation and Interpreting Studies are dominated during various periods by, let's say, one important paradigm, and after a certain time there is a more or less dramatic paradigm shift? Or would you describe it as different paradigms existing in parallel?*

B: The situation is rather that several paradigms are co-existing. But some of them are stronger than others. For the moment, for instance, the sociological aspect is very strong both in Translation and Interpreting Studies. So, of course, there are trends, and this may not be very surprising either. Translation Studies largely follows the development of its neighboring disciplines within linguistics and applied linguistics, and it has also looked further than that. So it's a natural development that you're inspired by new theories in other disciplines and that you explore what can be applied to your own discipline.

E: *And the driving force in other disciplines spreads to Translation and Interpreting Studies.*

B: Yes, and another driving force is senior researchers in a discipline or at a university department, since, naturally, their research interests will strongly influence students: they suggest topics that interest them and give presentations on their research areas, and supposedly they follow the development in other disciplines as well. So, of course, there are trends.

E: *Do you think that there has been a Kuhnian "scientific revolution" in Translation Studies? Not just in recent years, I mean has there been a major shift in Translation Studies at some point? Something that can easily be pin-pointed as having affected the whole field?*

B: One very important shift was Descriptive Translation Studies, which Gideon Toury first presented in the 1970s and the early 1980s. The whole 1990s was very much dominated by DTS and the drive to describe and analyze translations and their place in different cultures. And this can definitely be described as a revolution. But then that also depends on your personal interest; somebody else may define post-colonialism or some other areas in those terms. But these are not my areas, and I suppose that this applies to all disciplines. You cannot do everything, and in order to be good at something you have to accept that you pursue the path where you believe that you have some talent and something to add.

E: *You cover a very broad area. Many researchers in translation and interpreting – that I know of, anyway – I would perhaps not call them very specialized, but they are mostly dealing with either literary or technical translation or interpreting, and they often stay within the area where they are well-known and well-versed. But, for instance, you still work a lot with dialogue interpreting.*

B: Yes, it has turned out that way. I've had the privilege of working with interpreting and translation at so many different levels, in particular since I started here at TÖI in 1994. I have taught many courses, but I have also worked with the development of the training programs. And all the time I have combined it with research. The research I have done has gone hand in hand with the planning of training and my teaching. This is usually called research-related education, which is a beautiful word, but I really think I can say that in my case it is a firm belief that all training and education should be related to research. In my work at TÖI I have been privileged enough to see the links between research and training, and I have also been able to cross-fertilize the two areas in my work.

E*: I think I'm approaching the end of my questions. But I would like to ask you: what's your dream project? No limitations in time and money, anything?*

B: Right now, I'm very, very interested in the individual translator's style, but I'm not sure it's my dream project. There are so many things that are interesting. Another thing that I would like to explore is how secondary school students translate, I think it would be very exciting to give them an authentic translation commission, and give them the same conditions as a professional translator would get and then see what happens. They could for instance be commissioned with something that would be published on their school's webpage. Because in my latest major study, I saw that the language students who didn't have any professional experience of translation changed their way of thinking in the course of the translation task. They started thinking as the professional translators. Not as fast or as systematic of course, but they started taking the same factors into account – they started thinking in terms of usage of the text, audience and so forth, and their text changed accordingly. And I found those results fascinating, and it would be interesting to pursue them. But one of the challenges with the way we work is the amount of data. Our methods usually generate an excess of data, and if there's one thing that tires me now, it's dealing with all this data. I have traveled quite a lot with a tape or video recorder and a lot of tapes, and have recorded both interpreters and translators. But on the other hand, if you get the opportunity to work together with for instance PhD students, then you can share that burden.

E: *Finally what do you wish for Translation Studies?*

B: Oh, what do I wish? Do I have to wish? It's difficult.

E: *You may have visions.*

B: Oh, I wish that … I don't know what I would wish, but I am very worried about what I see as divisive tendencies in modern Translation Studies. This doesn't necessarily concern the scientific aspects, but it may also be channeled through the scientific aspects. Differences are stressed, different paradigm stand against each other. Also on the associative level, where various organizations seem to work very much for what seems to be the same goals. Maybe if you're really very well acquainted with the issues you see a difference, but for somebody seeing it from the outside it's difficult to understand why there are different organizations. And to me that's worrying. I know that on the other hand it is perhaps a sign of health, but I believe that since we aren't even a discipline in most countries and have to fight for, maybe not our raison d'être, but for our research culture, so to speak, I don't think we can afford to divide our forces, we need to gather them. So, I am a bit worried over the future.

E: *I may be putting words in your mouth here, but if you wish for something, is it mainly that there should be a more unifying force, creating a common ground of shared values instead of seeking to divide the field?*

B: Yes, although this might not be so popular today when interdisciplinarity is so popular and when Translation Studies is portrayed as a multidisciplinary and interdisciplinary field. And that's all fine, and I'm not the right person to criticize those

who seek theories and methods outside of Translation Studies – I've done that myself, after all – but I think sometimes this goes a bit too far. We aren't even aware of what has been done earlier in Translation Studies and what we can build on today, before we start looking for new theories that can be introduced to Translation Studies.

E: *You mean that we shouldn't overlook that there is a strong foundation that can be used to grow from?*

B: Yes, exactly.

E: *Well, I have come to the end of my questions, thank you so much for taking your time.*

B: Oh, I think I ended on a very negative note.

E: *No, on the contrary!*

Publications by Birgitta Englund Dimitrova*

Monographs, articles in scientific journals and articles with peer review in book volumes, other scientific works, and edited volumes

1977. *Yes/no-Questions in Bulgarian and Macedonian. Form.* Stockholm: Almqvist & Wiksell.

1979. With Jüri Waldin. "Sekundärspråksinterferens i primärspråket (på grundval av material från jugoslaviska invandrare i Sverige)." In *Kontrastiv lingvistik och sekundärspråksforskning*, B. Hammarberg (ed.), 19–29. Stockholm: Department of Linguistics, Stockholm University. [English translation of the title: Second language interference in the first language (on the basis of material from Yugoslavian immigrants in Sweden).]

1979. "The functions of yes/no-questions (on the basis of Bulgarian and Macedonian material)." In *Papers on Slavonic Linguistics* 1, 16–26 [Meddelanden från Institutionen för slaviska och baltiska språk 18]. Stockholm: Department of Slavic and Baltic Languages, Stockholm University.

1979. Editor, with B. Jangfeldt, L. Kleberg and B. Nilsson, of *Papers on Slavonic Linguistics, 1–3. Papers presented at the first Polish-Swedish Slavists' Conference at Mogilany, 1978* [Meddelanden från institutionen för slaviska och baltiska språk 18–20]. Stockholm: Department of Slavic and Baltic Languages, Stockholm University.

1981. "Za saotnošenieto meždu estestven pol i gramatičeski rod v sâvremennija bâlgarski ezik." In *Bâlgaristični izsledvanija*, 83–97. Sofia: University of Sofia. [English translation of the title: On the relationship between natural and grammatical gender in modern Bulgarian].

1983. "Časticata ‚nali' vâv vâprositelni izrečenija." In *Pârvi meždunaroden kongres po bâlgaristika. Sofija 23.5–3.6 1981 g. Dokladi. Istoričeski razvoj na bâlgarskija ezik. 2. Sâvremenen bâlgarski ezik,* 225–232. Sofia: BAN. [English translation of the title: The particle 'nali' in interrogative clauses.]

1984. "Časticata 'nali' vâv nevâprositelni izrečenija (funkcionalna charakteristika)." *Sâvremenna Bâlgarija. Tom 5,* 72–77. Sofia: BAN. [English translation of the title: The particle 'nali' in non-interrogative clauses.]

1986. "Za taka narečenata modalna upotreba na segašno vreme vmesto bâdešte." In *Proceedings of the Second Bulgaro-Scandinavian Symposium, held at Kungälv, 9–15 August, 1982* [Uppsala Slavic Papers 13: 2], 45–51. Uppsala: Department of Slavic, University of Uppsala [English translation of the title: On the so-called modal use of present tense instead of future tense].

1987. "Chajtov na švedski – njakoi problemi." In *Problems in connection with translation from Bulgarian into Swedish* [Gusli 2], 41–49. Gothenburg: Department of Slavic, University of Gothenburg. [English translation of the title: Hajtov in Swedish – a few problems.]

---

* Note: Publications 1977–1981 were published under the name Birgitta Englund.

1988. "Njakoi osobenosti na bâlgarskata gramatika s ogled prepodavaneto ì na švedi (upotrebata na osnovnite minali vremena v bâlgarski i švedski)." In *V pomošt na prepodavaneto po bâlgarski ezik na čuždestranni studenti*, 42–51. Sofia: St. Clement of Ohrid University of Sofia. [English translation of the title: Some special features in Bulgarian grammar particular for Swedes acquiring Bulgarian (use of the basic past tenses in Bulgarian and Swedish).]

1988. Review of *Lehrbuch der makedonischen Sprache* (V. Bojic and W. Oschlies, 1986. Munich: Sagner). *Slovo* 36: 85–90. Uppsala: Department of Slavic, University of Uppsala.

1989. "Tolkning". In *Invandrare och minoriteter* 4–5: 88–89. [English translation of the title: Interpreting.]

1991. "Att arbeta med tolk". In *Flyktingar och invandrare i sjukvården. Medicinska och kulturella aspekter. Läkardagarna i Örebro 1991. Svensk Medicin, utg. av Svenska Läkaresällskapet och Spri* 25: 287–297. [English translation of the title: Working with interpreters.]

1991. *När två samtalar genom en tredje. Interaktion och icke-verbal kommunikation i medicinska möten med tolk* [Rapporter om tvåspråkighet 7]. Stockholm: Centre for Bilingual Research, Stockholm University. [English translation of the title: When two people talk through a third one. Interaction and non-verbal communication in interpreted medical encounters.]

1991. "De icke-verbala komponenterna i tolkade samtal." In *Att tala utan ord. Människans icke-verbala uttrycksformer*, 51–62. [Kungl. Vitterhets Historie och Antikvitets Akademien, Konferenser 24]. Stockholm: Almqvist & Wiksell International. [English translation of the title: Non-verbal components in interpreter-mediated events.]

1993. "Metoder i empirisk forskning om översättning och tolkning." In *Studier i modern språkvetenskap* 10: 33–44. [Stockholm Studies in New Philology. N.S. 10]. Stockholm: Almqvist & Wiksell International. [English translation of the title: Methods in empirical research on translation and interpreting.]

1993. "Semantic change in translation – a cognitive perspective." In *Translation and Knowledge. SSOTT IV,* Y. Gambier and J. Tommola (eds), 285–296. Turku: University of Turku, Centre for Translation and Interpreting.

1993. "Turntaking in interpreted discourse: interplay of implicit and explicit rules." In *Problemi na sociolingvistikata. Ezikovata situacija v mikro- i makrosocialnite obštnosti*, K. Cankov (ed.), 15–21. Veliko Târnovo: St. Cyril and St. Methodius University of Veliko Tarnovo.

1994. "Statistical analysis of translations (on the basis of translations from and to Bulgarian, Russian and Swedish)." In *Scandinavian Working Papers on Bilingualism* 9: 87–102. [Available at: http://www2.tolk.su.se/Birgitta_Englund_Dimitrova/documents/statistical_analysis_of_translations.1994.pdf.]

1995. "Omissions in consecutive interpretation: evidence of a decision component." In *Podobaet' pamjat' s"tvoriti. Essays to the Memory of Anders Sjöberg* [Stockholm Slavic Studies 24], P. Ambrosiani, B. Nilsson and L. Steensland (eds), 73–82. Stockholm: Almqvist & Wiksell International.

1995. "Examensarbetet på översättarutbildningen – en akademisk hybridgenre?" In *Konsten att skriva*, 27–28 [Pu-rapport 8]. Stockholm: Stockholm University, Centre for Educational Development. [English translation of the title: The dissertation at the translation programme – an academic hybrid genre?]

1995. "Teori och praktik i universitetens översättarutbildningar." In *Språkundervisning på universitet. Rapport från ASLA:s höstsymposium i Göteborg, 11–13 november 1993* [ASLA:s skriftserie 7], L.-G. Andersson and F. Börjesson (eds), 50–57. Uppsala: ASLA, Svenska

föreningen för tillämpad språkvetenskap. [English translation of the title: Theory and practice in university translation programmes.]

1996. "New methods in translation research – new horizons in translation studies?" In *Proceedings of the XIV World Congress of the Fédération Internationale des Traducteurs, vol. 2*, 856–864. Melbourne: AUSIT.

1996. "Översättningsförmåga och översättarkompetens." In *ASLA-information* 22 (3): 37–45. [English translation of the title: Translation ability and translator competence.]

1996. "Förståelseprocesser vid översättning." In *Språkförståelse. Rapport från ASLA:s höstsymposium i Lund, 9–11 november 1995*, G. Håkansson and U. Nettelbladt (eds), 17–26. [ASLA:s skriftserie 9]. Uppsala: ASLA, Svenska föreningen för tillämpad språkvetenskap. [English translation of the title: Comprehension processes in translation.]

1996. Editor of *Temanummer om forskning inom tillämpad lingvistik vid Stockholms universitet. ASLA-information* 22: 3. [English translation of the title: Special issue on research in applied linguistics at Stockholm University.]

1996. *Verksamhetskunskap och forskningskompetens. Forsknings- och utvecklingsprogram för Tolk- och översättarinstitutet* [TÖI rapport 5]. Stockholm: Stockholm University, Institute for Interpretation and Translation Studies. [English translation of the title: Activity knowledge and research competence. Research and development programme at the Institute for Interpretation and Translation Studies.]

1997. "Utbildning av lärare till högskolans översättarutbildningar – rapport från ett pågående försök." In *Kvalitet genom decentralisering – universitetet i centrum* [Pu-rapport 4], C. Sjöström (ed.), 67–69. Stockholm: Stockholm University. [English translation of the title: Training teachers for university translation programmes – report from an ongoing experiment.]

1997. "Translation of dialect in fictional prose. Vilhelm Moberg in Russian and English as a case in point." In *Norm, Variation and Change in Language* [Stockholm Studies in Modern Philology. N.S. 11], J. Falk (ed.), 49–65. Stockholm: Almqvist & Wiksell International.

1997. "Degree of interpreter responsibility in the interaction process in community interpreting." In *The Critical Link. Interpreters in the Community. Papers from the First International Conference on Interpreting in Legal, Health, and Social Service Settings, Geneva Park, Canada, June 1–4, 1995* [Benjamins Translation Library 19], S. E. Carr, R. Roberts, A. Dufour and D. Steyn (eds), 147–164. Amsterdam/Philadelphia: John Benjamins.

1997. "Translation ability and translatorial competence." In *Translation and Interpreting*, B. Svane (ed.), 31–32. Stockholm: Stockholm University.

1998. "Tradition och förnyelse i metoder inom översättningsforskningen – exempel på analys av think-aloud-protocols." In *Meddelanden från Institutionen för nordiska språk vid Uleåborgs universitet, Serie B Nr 23*, I. Sorvali (ed.), 7–30. [English translation of the title: Tradition and renewal of research methods in translation studies: An example of an analysis of think-aloud protocols.]

1998. "Bra opposition – vad är det och hur får vi det?" In *Hur förbättrar vi forskarutbildningen* [Pu-rapport 4], G. Sebardt (ed.), 60–62. Stockholm: Stockholm University. [English translation of the title: Good feedback: What is it and how do we get it?]

1999. With Kerstin Jonasson. "Translation Ability and Translatorial Competence: Expert and Novice Use of Dictionaries". Paper presented at the Symposium on Translation and Interpreting, organized by Stockholm University in cooperation with Humboldt University, Berlin, in Stockholm, January 1999. [Available at http://www2.tolk.su.se/Birgitta_Englund_Dimitrova/documents/englund_dimitrova_jonasson.1999.pdf.]

1999. "Tolkens roll: perspektiv från nyare forskning och implikationer för tolkutbildningen." In *Svensk-ryskt tolkundervisningssymposium, Moskva, MGU 24–26 mars 1999.* Moscow: MGU, Department of Germanic Languages. [English translation of the title: The interpreter's role: Perspectives from recent research and implications for the education of interpreters.]

2000. "Accreditation." (Summary of discussion at an on-line symposium on Innovation in Translator and Interpreter Training.) *Across Languages and Cultures* 1 (2): 247–249.

2000. With Kenneth Hyltenstam. "Preface." In *Language Processing and Simultaneous Interpreting. Interdisciplinary Perspectives*, B. Englund Dimitrova and K. Hyltenstam (eds), ix–xv. Amsterdam/Philadelphia: John Benjamins.

2000. Editor, with K. Hyltenstam, of *Language Processing and Simultaneous Interpreting: Interdisciplinary Perspectives* [Benjamins Translation Library 40]. Amsterdam/Philadelphia: John Benjamins.

2000. Editor of *Översättning och tolkning. Rapport från ASLA:s höstsymposium Stockholm 5–6 november 1998* [ASLA:s skriftserie 12] Uppsala: ASLA. [English translation of the title: Translation and interpreting.]

2001. "En smålänning i engelsk och fransk skepnad. Om översättning av dialekt i skönlitteratur." *Folkmålsstudier. Meddelanden från Föreningen för nordisk filologi* 40: 9–27 [English translation of the title: A Swede from Småland in English and French guise. On the translation of dialect in fiction].

2001. "Egennamn vid översättning från ryska till svenska. Olika problem för olika översättare." In *Langage et réference. Mélanges offerts à Kerstin Jonasson à l'occasion des ses soixante ans* [Studia Romanistica Upsaliensia 63], H. Kronning, C. Norén, B. Novén, G. Ransbo, L.-G. Sundell and B. Svane (eds), 129–136. Uppsala: Acta Universitatis Upsaliensis. [English translation of the title: Translating proper names from Russian into Swedish. Different problems for different translators]

2001. "Training the interpreter trainers – the Swedish experience". Paper presented at the conference Critical Link 3, Montreal, May 2001. [Available at http://www2.tolk.su.se/Birgitta_Englund_Dimitrova/documents/paper_montreal_2001.pdf.]

2002. "Training and educating the trainers. A key issue in translators' training." In *Teaching Translation and Interpreting 4. Building Bridges* [Benjamins Translation Library 42], E. Hung (ed.), 73–82. Amsterdam/ Philadelphia: Benjamins.

2002. Review of *Getting started in interpreting research* (D. Gile, H.V. Dam, F. Dubslaff, B. Martinsen and A. Schjoldager (eds), 2001. Amsterdam/Philadelphia: John Benjamins). *Across Languages and Cultures* 3 (2): 267–270.

2002. Review of *The Interpreting Studies Reader* (F. Pöchhacker and M. Shlesinger (eds), 2002, London: Routledge). *Across Languages and Cultures* 3 (2): 267–270.

2002. "Från sydöstra Sverige till sydöstra Europa – vilken bulgariska talar Emil i Lönneberga?" In *Explorare necesse est. Hyllingsskrift till Barbro Nilsson* [Acta Universitatis Stockholmiensis, Stockholm Slavic Studies 28], P. Ambrosiani, E. Löfstrand, L. Nordquist and E. Teodorowicz-Hellman (eds), 81–89. Stockholm: Almqvist & Wiksell International. [English translation of the title: From the South-east of Sweden to the South-east of Europe: What Bulgarian does Emil from Lönneberga speak?]

2003. "Explicitation in Russian-Swedish translation: Pragmatic and sociolinguistic aspects." In *Swedish Contributions to the 13th Congress of Slavicists, Ljubljana August 2003* [Slavica Lundensia supplementa 2], B. Englund Dimitrova and A. Pereswetoff-Morath (eds), 21–31. Lund: Department of Central European Studies, Lund University. [Available at: http://www2.tolk.su.se/Birgitta_Englund_Dimitrova/documents/englunddim.explic.pdf.]

2003. Editor, with A. Pereswetoff-Morath, of *Swedish Contributions to the 13th Congress of Slavicists, Ljubljana August 2003* [Slavica Lundensia supplementa 2]. Lund: Department of Central European Studies, Lund University.

2004. "Orality, literacy, reproduction of discourse and the translation of dialect." In *Dialektübersetzung und Dialekte im Multimedia*, I. Helin (ed.), 121–139. Frankfurt am Main: Peter Lang.

2005. "Literal translation in the translation process of professional translators." In *New Tendencies in Translation Studies. Selected Papers from a Workshop, Göteborg 12 December 2003* [Gothenburg Studies in English 90], K. Aijmer and C. Alvstad (eds), 29–39. Gothenburg: University of Gothenburg, Department of English.

2005. *Expertise and Explicitation in the Translation Process* [Benjamins Translation Library 64]. Amsterdam/Philadelphia: John Benjamins.

2005. "Combining product and process analysis. Explicitation as a case in point." In *Empirische Untersuchungen zum Übersetzungs- und Dolmetschprozess,* A. Künzli (ed.), *Bulletin Suisse de linguistique appliquée* 81: 25–39.

2006. Editor, with H. Landqvist, of *Svenska som källspråk och målspråk: aspekter på översättningsvetenskap. Artiklar från en forskarkurs vid Göteborgs universitet höstterminen 2005* [Översättningsstudier vid Göteborgs universitet 9]. Gothenburg: University of Gothenburg. [English translation of the title: Swedish as source and target language: Aspects of translation studies. Papers from a doctoral school at the University of Gothenburg, autumn semester 2005.]

2007. Review of *New Trends in Translation Studies. In Honour of Kinga Klaudy* (K. Károly and Á. Fóris (eds), 2005, Akademiai Kiado). *Across Languages and Cultures* 8 (1): 137–142. [Available at http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-10415.]

2007. "Tjugo år av svensk översättningsvetenskap." In *Gränslösa texter. Perspektiv på översättning*, Y. Lindqvist (ed.), 13–31. Uppsala: Hallgren & Fallgren. [English translation of the title: Twenty years of Swedish translation studies.]

2007. Editor, with C. Wadensjö and A.-L. Nilsson, of *The Critical Link 4. Professionalisation of Interpreting in the Community. Selected Papers from the 4th International Conference on Interpreting in Legal, Health and Social Service Settings, Stockholm, Sweden, 20–23 May 2004* [Benjamins Translation Library 70]. Amsterdam/Philadelphia: John Benjamins.

2008. "Svensk översättningsvetenskap från 1986 och framåt. Lista över publikationer i bokform eller som elektronisk resurs". [English translation of title: Swedish translation studies from 1986 and onwards. List of books and electronic resources.] [Available at http://www2.tolk.su.se/Birgitta_Englund_Dimitrova/documents/sv._versvetenskap.feb.08.pdf.]

2009. With Elisabet Tiselius. "Exploring retrospection as a research method for studying the translation process and the interpreting process." In *Methodology, Technology and Innovation in Translation Process Research. A Tribute to Arnt Lykke Jakobsen*, I. M. Mees, F. Alves and S. Göpferich (eds), 109–134. Copenhagen: Samfundslitteratur.

2010. Review of *Interpreting Studies and Beyond. A Tribute to Miriam Shlesinger* (F. Pöchhacker, A. Lykke Jakobsen and I. Mees (eds), 2007. Copenhagen: Samfundslitteratur) and Review of *Efforts and Models in Interpreting and Translation Research: A Tribute to Daniel Gile* (G. Hansen, A. Chesterman and H. Gerzymisch-Arbogast (eds), 2008, Amsterdam/Philadelphia: John Benjamins). *Interpreting* 12 (2): 249–258. [Available at http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-42375.]

2010. "The translation process". In *Handbook in Translation Studies*, Y. Gambier and L. Van Doorslaer (eds), 406–411. Amsterdam/Philadelphia: John Benjamins. http://www.benjamins.com/online/hts/

Teaching material

1999. [2nd revised version 2002]. *Examensarbetet på översättarutbildningen. Några tips för studerande.* Stockholm: Institute for Interpretation and Translation Studies. [English translation of the title: The dissertation at the translation programme. Some hints for students.] [Available at http://www.tolk.su.se/content/1/c6/06/41/10/exarbete.PDF].

Popular works

1977. Swedish preface to *Under the Yoke*, Ivan Vazov (auth.), 9–10. Stockholm: Tidens förlag.
1980. "Nikolaj Chajtov och hans vilda berättelser." *Slavisk kulturrevy* 1: 29–31. [English translation of the title: Nikolaj Chajtov and his wild stories.]
1981. "Sultana och Irina – två kvinnobilder hos Talev och Dimov." *Artes* 1: 85–96. [English translation of the title: Sultana and Irina: Two portraits of women by Talev and Dimov.]
1982. "Sluge Peter – en hjälte i bulgarisk folkdiktning." *Slavisk kulturrevy* 3: 6–7. [English translation of the title: Cunning Peter – a hero in Bulgarian folk poetry.]
1995. "Simultantolkning." In *Nationalencyklopedin, band 16*. [English translation of the title: Simultaneous interpreting.]
1995. "Tolkning; TV-översättning." In *Nationalencyklopedin, band 18*. [English translation of the title: Interpreting; TV translation.]
2001. "Är det skillnad på tolk och tolk?" In *Språkbitar*, J. Nystedt (ed.), 34–35. Stockholm: Svenska förlaget. [English translation of the title: Are all interpreters the same?]
2001. "Emil i Lönneberga – en läspande fransman?" In *Språkbitar*, J. Nystedt (ed.), 97–98. Stockholm: Svenska förlaget. [English translation of the title: Emil of Lönneberga – a lisping Frenchman?]

Notes on contributors

Ruth Almog received her Master's Degree in Translation and Interpreting Studies from the Bar Ilan University, Israel. She is a practicing translator and teacher of EFL. She has been teaching Pedagogic Translation Studies at Mevo'ot Eron High School since 1998. Her master's degree centered mainly on meta-linguistic awareness among high-school students who have taken the Pedagogic Translation Studies Course.

Cecilia Alvstad is Associate Professor of Spanish at the University of Oslo. In 2011 she is Visiting Researcher at The Institute for Interpretation and Translation Studies at Stockholm University. Alvstad's current research interests concern translational relations between Scandinavia and the Americas as manifested in literary translations, migrant letters and travel writing. Alvstad has also conducted research in the field of translation of children's literature. She is the vice president of the European Society for Translation Studies.

Claudia V. Angelelli is Professor of Spanish Linguistics at San Diego State University. She is the author of *Medical Interpreting and Cross-cultural Communication* (Cambridge University Press), and *Revisiting the Role of the Interpreter* (John Benjamins) and the co-editor of *Testing and Assessment in Translation and Interpreting Studies* (John Benjamins). She is President of ATISA, World Project Leader for ISO Standards on Community Interpreting and Director of The Consortium of Distinguished Language Centers.

Erik Angelone is currently Assistant Professor of Translation Studies at Kent State University. He received his PhD in Translation Studies from the University of Heidelberg and his M.A. in Intercultural Communication from the University of Maryland Baltimore County. His research interests include process-oriented translator training, intercultural communication, and corpus-based translation studies. Angelone recently co-edited *Translation and Cognition* (John Benjamins) with Gregory Shreve.

Andrew Chesterman has recently retired from his post as Professor of Multilingual Communication, department of Modern Languages, University of Helsinki. He has been based in Helsinki for about forty years, working in contrastive analysis and translation studies. Chesterman's main research interests are research methodology and translation theory (e.g. norms, universals, ethics, sociological turn, translator training).

Barbara Dragsted is associate professor at the Department of International Language Studies and Computational Linguistics, Copenhagen Business School. She teaches specialized communication and translation and is involved in various research projects under the Center for Research and Innovation in Translation and

Translation Technology (CRITT). Her research interests include cognitive processes in translation, translation technology, LSP translation, terminology and knowledge sharing, as well as translation and interpretation hybrids.

Maureen Ehrensberger-Dow is a Canadian psycholinguist who has been involved in research into multilingualism and translation in Switzerland for the past 10 years. From a learner-based perspective driven by an interest in second language acquisition, she has become increasingly drawn into research into language practitioners' strategies and expert knowledge. She is currently principal investigator of the Capturing Translation Processes research project. She is affiliated to the Zurich University of Applied Sciences (ZHAW), Winterthur, Switzerland.

Daniel Gile, professor at ESIT (Université Paris 3 Sorbonne Nouvelle) is a former mathematician turned scientific and technical translator, then conference interpreter, then translator and interpreter trainer and researcher. He holds a PhD in Japanese and a PhD in linguistics and has published more than 200 papers and 3 books and co-edited 7 collective volumes on translation, interpreting and research. He is a former president of the European Society for Translation Studies, and his interests include the cognition of conference interpreting, translator and interpreter training, research training and translation and interpreting research as an object of study.

Inge Gorm Hansen is Associate Professor in the Department of International Language Studies and Computational Linguistics at the Copenhagen Business School. She teaches interpreting and intercultural communication and is practicing translator and interpreter. Her research interests include interpreting, translation process studies and translation technology, terminology and knowledge sharing, and translation and interpretation hybrids.

Rob Hartsuiker obtained a master's degree in experimental psychology from the University of Leiden and a PhD from the University of Nijmegen. After holding postdoctoral research positions in Nijmegen and Edinburgh, he moved to Ghent University where he is now professor of psychology. He investigates the cognitive mechanisms of language processing, in particular speech errors and self-repair, sentence production, the neurocognition of language, and bilingualism.

Adelina Hild received her master's degree in general and English linguistics and her PhD in applied linguistics at the University of Cambridge, UK. She followed an MA course in translation and interpreting from St. Kliment Ohridski, Bulgaria. After the completion of her doctorate she was awarded a postdoctoral fellowship at centre for Treatment of Multilingual Information (ETI, Switzerland), and followed by a position as a research associate with CRIT, State University of New York (USA). She is working as a freelance professional interpreter, interpreter trainer and intercultural communication consultant. Her research interests focus on discourse processing and expertise psychology, cognitive accounts of expertise in simultaneous interpreting and multilingualism.

Arnt Lykke Jakobsen is professor of translation and translation technology at Copenhagen Business School and director of the Centre for Research and Innovation in Translation and Translation Technology (CRITT). After developing the Translog

software, his main research effort has been to develop a methodology for translation process research based on keylogging and eyetracking.

Gard Jenset received his PhD from the University of Bergen. He is associate professor of English linguistics at the Faculty of Education, Bergen University College, Norway, and his research interests span language learning and language processing, historical corpus linguistics and language change, and quantitative methods in language research.

Alexander Künzli has a Master's degree in translation and psychology from Geneva University and a PhD in French Linguistics from Stockholm University. He is a lecturer at the Institute of Translation and Interpreting at Zurich University of Applied Sciences. Current research interests include translation revision, audiovisual translation, and the translation of Swedish crime fiction.

Christian Lachaud is a Cognitive Psychologist trained in Psycholinguistics and Cognitive Linguistics (France, Switzerland, USA, Norway) at the Department of Literature, Area Studies (ILOS), and European Languages at the University of Oslo, Norway. His research interests focus on word recognition in spoken and signed languages, the structure of the bilingual mental lexicon, translation, and conceptual blending in figurative meaning processing. He bases his work on advanced techniques for investigating language behaviors and their correlated brain activity (Eye-tracking, EEG Power and Coherence analysis, Statistical modeling, Connexionist Simulation).

Isabel Lacruz is Assistant Professor of Spanish and Translation Studies in the Institute for Applied Linguistics of the Department of Modern and Classical Language Studies at Kent State University. She received her PhD in Experimental Psychology from Kent State University in 2005. Her research interests include the psycholinguistic aspects of translation, bilingual memory, and second language acquisition. Her most recent research involves the use of eye tracking methodologies to study the translation process.

Brian Mossop has been a translator, reviser and trainer at the Canadian federal government's Translation Bureau for the past 37 years and a course director at the York University School of Translation in Toronto for the past 32 years. He is the author of the textbook *Revising and Editing for Translators* (St Jerome) as well as some 40 articles on translation theory and translation pedagogy.

Ulf Norberg holds a PhD in German linguistics from Uppsala University, Sweden. He now lectures translation theory and practical translation at the Institute for Interpretation and Translation Studies, Stockholm University. His dissertation *Übersetzen mit doppeltem Skopos. Eine empirische Prozess und Produktstudie* (2003), in the field of empirical translation process research, has been followed by articles on paratexts of translation and the translation of dialect.

PACTE has been a competitive research group at the Universitat Autònoma de Barcelona, Spain, since 1997. PACTE group's main research interests are: Empirical and experimental-based research on translation competence and its acquisition in written translation; Translator training; Empirical and experimental research in Translation Studies; The use of new technologies in translation research. The Group

was awarded the Universitat Autònoma de Barcelona Prize for Outstanding Research (PREI2010 UAB) for the article "Results of the Validation of the PACTE Translation Competence Model: Acceptability and Decision Making" (*Across Languages and Cultures* 10/2, 2009).

Antin Fougner Rydning is Professor of French and Translation studies in the Department of Literature, Area Studies (ILOS), and European Languages at the University of Oslo, Norway. She has been a visiting professor at Ecole Supérieure d'Interprètes et de Traducteurs in Paris and has served as Head of the French department and Head of Research at ILOS. She has been partner in the EU Commission research project EYE-to-IT (6FrP). She has been chairman of The State Authorized Translators' Guild in Norway, and has worked as a state authorized translator and conference interpreter in France and Norway for more than 30 years. Her major research interests include translation theory, cognitive processes in translation, and empirical methods in translation research.

Miriam Shlesinger is a practicing translator and interpreter. She has been teaching translation and interpreting (both theory and practice) at Bar Ilan University since 1978. Her doctoral research centered on cognitive processes – particularly, attention and working memory – in simultaneous interpreting. She is the co-editor (with Franz Pöchhacker) of the *Interpreting Studies Reader* (Routledge) and, since 2004, of the journal *Interpreting* (John Benjamins), as well as Associate Editor of the *Benjamins Translation Library*. She is affiliated to the Department of Translation and Interpreting Studies, Bar Ilan University, Israel and the Department of Language Management and Language Practice, University of the Free State, RSA.

Gregory M. Shreve is Emeritus Professor of Translation Studies at Kent State University in the Institute for Applied Linguistics (IAL) of the Department of Modern and Classical Language Studies. He received his doctorate in Anthropology from Ohio State University in 1975. Shreve is the Founding Director of the IAL, home of one of America's leading graduate professional translator training programs. His research interests include computer-assisted translation, cognitive processes in translation, and empirical methods in translation research.

Šárka Timarová works as a research assistant at Lessius University College (Belgium). She trained as a conference interpreter at Charles University (Czech Republic), and followed her studies with a postgraduate research training programme (DEA in Interpreting) and Continuing Education Certificate for Interpreter Trainers (EP bursary) at ETI (University of Geneva, Switzerland). She is currently working on her doctoral dissertation (Charles University, Czech Republic), focusing on the involvement of working memory in simultaneous interpreting and the applicability of these findings to interpreter training.

Elisabet Tiselius holds an M.A in Conference Interpreting. She is currently a PhD student at the Department of Foreign Languages, University of Bergen, Norway and the Centre for Research on Bilingualism, Stockholm University, Sweden. Her PhD thesis explores the development of expertise in conference interpreting. She is

also active as conference interpreter accredited to the European institutions, and as community interpreter certified between Swedish and English.

Sonia Vandepitte is a lecturer at University College Ghent Translation Studies, where she teaches translation studies, English and translation into and from Dutch. She also coaches master student translation offices. She has published on intonation, causal expressions, knowledge retrieval, methodology in translation studies, translation competences, anticipation in interpreting and international translation teaching projects. She supervises research projects on translation corpora, a Dutch-English clarity tool and electronic feedback in language teaching.

Gun-Viol Vik-Tuovinen is a Senior lecturer, University of Vaasa, Finland. Since 1987 she has held a full-time teaching and research position in the area of interpreting and translation at the University of Vaasa. Her doctoral dissertation from 2006 was about interpreting on different levels of professionalism. She has published about thirty scientific papers, mainly about interpreting. She is a freelance conference interpreter between Scandinavian languages and Finnish.

Cecilia Wadensjö is Professor of Interpreting and Translation Studies at Stockholm University, Sweden. She has published extensively on interpreting as a linguistic and social phenomenon, most significantly the monograph *Interpreting as Interaction* (Addison Weasley Longman) and the co-edited volume *The Critical Link 4: Professionalization of Interpreting in the Community* (John Benjamins). She also holds a University diploma in Interpreting and is Certified Interpreter between Swedish and Russian.

Index

A
abstract nouns 72
academic field 13–14, 17, 19
acceptability (of translations) 70, 81–82, 204, 208, 227, 295, 317, 320–325, 329–330, 332, 334, 336–341
access to expert knowledge 270
acquisition of translation competence 318
additions 57, 60, 123
adequacy 75, 81–82, 252
adversarial humor 219, 229
agency 233, 240, 244
agent prototypicality 67, 77–78
ambiguity 17, 61, 74, 152, 176
amount of embedding 249, 253, 257, 259, 261
anchor word 46, 48, 50, 52–53
anecdotal evidence 69–70
annotated protocol 100–102, 107–108, 116
answering odds 180
applied research 19
Argument Selection Principle 71
assessment 274, 284–286, 291, 293–294
audience response 187–188, 193, 196
audiovisual media 198–199
audiovisual translation 188, 199
authentic texts 151

B
basic research 14, 19
behavioural data 38
behavioural pattern 182
between-subject 73, 78
bilingual sub-competence 319
bilingual word recognition 82, 85
biplot 277–278, 280–281, 284
blend 173
bodily experience 176
borrowings 152
bottom-up approach 81
bridge the gap 3, 16
British English 74, 172, 254

C
Carroll's scales 269, 274, 285
causation 71
Celex database 74
Chi-square test 78
chronic illnesses 231–232, 237, 239–242
classroom intervention 153
clause complement 257–258, 261–262
clear goals 270
COBUILD corpus 74
coded file 75
cognate avoidance 113–114, 118
cognates 69, 74, 112–114, 152, 161–162, 207
cognitive effort 93, 95, 98, 103–105, 107, 110, 115, 117, 127, 144, 169, 170, 188–189
cognitive process(es) 170, 219–220, 228–229, 274, 325–326, 328–329, 334
cognitive processing 29, 95, 105, 108, 121–123, 127, 169–170, 177, 251
coherence 264
collaborative learning 155
collocation 62, 152, 157–158
complex conceptual metaphors 170
complexity metrics 249–250, 252–253, 256, 258, 263, 265
comprehension 1, 3, 37–38, 40–41, 45, 47–48, 93–94, 96, 135, 144, 174, 251–252, 257, 263–265, 272, 278–279, 295, 300, 327–328, 333–335
comprehension processes 37–38, 41
computational resources 251, 258
conceptual architecture 174
conceptual clarity 170, 174
conceptualization 174
conceptual level 171
conceptual metaphors 170
concrete nouns 72, 227
conditional knowledge 339
congruence 24–25
connectors 57, 63
connotation 152, 197
consent form 76, 192
contrastive analysis 24–25, 151–152, 249
contrastive syntactic-semantic structures 69
controlled psycholinguistic investigation 67, 69
Correspondence Analysis 277
correspondence 24–27, 124–126, 175, 177, 258
creativity phases 223
creativity process 221, 226
creativity techniques 221
critical discourse analysis 231–232, 237–238, 245–246
cross-domain associations 171
cross-linguistic differences 67–68
cultural knowledge 19, 310
culture-specific items 152

D
decision-making 320–323, 326, 328, 330, 341
declarative knowledge 319, 339
delayed repair 115–116
deliberate practice 270, 276
deliteralization 27, 33–34
dependent variables 95, 175, 320
developmental level 249, 252, 255, 257, 259, 267
deverbalization 29–30
Dialang test 179
dialect 76, 85, 152, 188
directionality 317, 339
direct translation 322, 324–325, 327, 331–333, 335–339, 342
drafting phase 40

E
ear-voice span 121, 123, 136–138, 140, 143
ecological validity 70, 81, 202, 254, 353
efficacy of the translation process 320, 323, 341
Effort Models 204–205
elicitation 163
embedding type 249, 253, 258, 264
embodied meaning 174
empirical-experimental research 317
empirical translation 219
end revision 40
English-Dutch translation 67
environmental characteristics 172
E-prime script 178
error probability 175
errors 201, 204–213
evidence-based practice 13, 16, 19
experience 252–253, 260–262, 269–272, 274–276, 278–292, 298–299, 305–306, 308, 310–312, 318, 320–322, 346–347, 350–351, 358
experiential grounding 174
expert 249, 252–253, 260, 264–266, 270, 272, 275, 282, 292–293, 305, 318, 326, 334, 351
expertise 252–253, 256, 258–259, 264, 266–267, 269–276, 291–294, 313–314, 320–321, 324, 349, 351
expertise in translation 219, 320–321, 324
Expertise project 40
expert knowledge 305, 318, 326, 334
expert processing 249, 264
explanation task 154
explanatory power 23, 29, 33
explicative knowledge 339
explicitation 23, 29–31, 65, 190, 201, 285, 355
extralinguistic culture-bound references 190
extra-linguistic sub-competence 319
external observers 57–58, 65–66
eye-key span 121, 123, 128–134
eye movement(s) 38, 41, 45–47, 53, 70, 98, 197
Eye-to-IT project 37, 41–42, 128
eye-tracking 2, 36, 38, 45, 53, 65, 69, 95–96, 98, 101, 103–105, 107, 113, 116, 118, 121, 127–128, 131, 134, 144, 187–189, 192–193, 195–196, 198, 219, 352

F
face-to-face 231, 234–237, 239
fansubs 188–189
feedback 75, 155–156, 159–160, 165, 188, 226, 237, 270, 309
filled pauses 99–100, 102–103, 107, 109, 116–118, 125
filtering 180

G
gaze-to-word mapping (GWM) 42, 50–52, 127–129, 133–134
gender 125, 152, 162, 168, 262
genre 33, 152, 220, 324, 327, 334
gradient of complexity 184
grammar 157–158
graphemic word length 74
gravitational pull 29
gridding 152, 157–158, 163

H
health care provider/patient interaction 231, 246
Hebrew 149, 151, 157–163, 168
high school 149, 168, 347
homonymy 152
humor translation 220
hypothesis 23–24, 26–30, 32–34, 38, 40, 46, 104, 174

I
identification and solution of translation problems 317–318, 320–321, 323, 325–326, 329–330
idioms 28, 114, 156, 158, 162–163, 176, 211
independant variables 95, 175, 181, 256, 320
independent sample t-tests 74
infelicities 201, 204–206, 208–210, 212
information distribution pattern 75
information processing 93, 100, 106, 174, 250
informativeness 269, 274, 285–289, 291–292
informativity 256, 264
instrumental sub-competence 319–320
integrative processes 172
intelligibility 269, 274, 285–289, 291
intentions 57–59, 61–62, 64–66
interaction 180
interference 28–29, 69, 205, 210
interim solutions analysis 28
interpreted-communicative event 234
interpreter's role 231, 234, 236–237, 244–245, 304, 312–313
interpreter training 301, 303, 341
inter-stimulus interval 178

intralinguisitc culture-bound elements 190
inverse translation 322, 324–325, 327, 329, 331–339, 342
isolated sentence translation 81
isolated word translation 72, 81

K
keylogging 37–38, 53, 121, 131, 134, 144, 351–352
keystroke logging 28, 30, 40, 99, 127–128, 134, 144
knowledge about translation 319–320, 322–323

L
language history questionnaire 76, 80
language-pair 201, 213, 215
language production 58, 61, 66, 70, 94–95, 118, 205, 237
language production acts 64–65
language-related skills 149
language specific varieties 172
lateral thinking 221, 226, 229
lexical decision 69, 80
lexical parallelism 222, 227–228
lexical planning 95
lexical transfer disfluencies 110–111, 118
lexis 157, 327
linguistic competence module 67
linguistic complexity 249–253, 255–259, 263–264, 266
linguistic context 82, 85
linguistics 67, 174, 238, 355–356
literal translation 23–33, 71, 73, 124–125
long experience 269–270, 275–276, 279–284, 286–287, 289–292
longitudinal study 154, 318, 352

M
macrostrategies 264
main clause 257
management of resources 249
matriculation exam 149, 151, 154
meaning 29, 33, 53, 57, 60–64, 66, 69–70, 80, 97, 124–125, 128, 131–132, 144, 152, 154, 157–158, 163–164, 170–177, 179, 206, 208–212, 222–223, 225–228, 231–232, 242, 285, 295, 300, 307, 312, 324, 334
meaning construction 53, 170
meaning retrieval 185
mechanisms of comprehension 171
medical history 231–232
mental processing 174
meta-linguistic awareness 149–150, 153–154, 164–165
metaphorical expressions 172
metaphorization 170, 227
methodology 201
metonymy 67
missing values 180
monitoring 269, 271–272, 274, 277, 280–282, 290–292, 295, 298
monolingual processing 264
Morae software 75

N
naturalistic research 201
near-cognates 74
norm 2, 32–33, 63, 71, 152, 187–189, 196–199, 203, 206, 208, 240, 290, 301–304, 208–312
novice 249, 252–253, 260, 263–265, 270

O
omissions 201, 204–208, 210–215
operative knowledge 339
outstanding performance 270, 276, 351

P
paired set 74
paraphrasing 226, 228
participant and non-participant observations 234
pause(s) 37, 40–41, 43, 47, 69, 95–96, 98–114, 116–118, 123, 125, 127, 132, 136, 142, 203, 207, 219, 221, 325
PC class 75
pedagogical approaches 153
perception 4, 71, 136, 141–142, 155, 157, 159, 187, 189, 193–194, 243, 273, 278, 295, 307, 317, 319, 323, 330–332, 336, 338–339
performance accuracy 258–259, 264
phonological word length 74
polysemy 74, 152
post-interpretative process 251, 264
power differentials 231–233, 237–239, 243
practical knowledge 301, 305–306, 308–310, 315
predicate 68, 71–72, 74–75, 257
pre-test 70, 73–76, 79, 87, 89–90, 176–177
primary conceptual metaphors 170
priming 70, 72–73, 81–82
probe correspondence 175
probe words 177
problem-solving 154
procedural knowledge 319
process 45, 53, 69, 72, 127, 133, 144, 153–156, 159, 162, 168, 220, 223, 269–273, 284, 289, 291, 307, 320, 345, 349–353, 356
processing difference 174
processing difficulty 253, 257–258, 263–264
processing limitations 263
processing literal meaning 170
processing load 251, 256, 264
processing metaphorical meaning 170
processing problems 271–272, 274, 277–280, 282, 284, 290–292, 295
process research 1–3, 7, 45, 107, 116, 121–123, 126–127, 138, 219, 269, 345, 349–356
product 269–271, 273–274, 284–286, 289, 291, 295–297

production 1, 3, 37–41, 46–47, 57–58, 61, 64–66, 70, 91, 93–96, 98–99, 106–110, 112–113, 115, 117–118, 123, 128–133, 135, 144, 154, 205–206, 209–210, 251–252, 257, 272–274, 295, 325, 339
production processes 38, 41, 93–94, 99, 135, 272
professional field 14
professional judgement 13, 18, 20
professional training 152, 356
proficiency 164, 301–302, 306–307, 313
prototypical translation problems 326, 330, 333–334
psycholinguistics 67, 69–70
psycho-physiological components 318–319

Q
quality 273–274, 293–294
quality control 34
questionnaire 187–189, 196, 198, 302, 308–312, 315, 321–323, 330, 333–334, 336, 339, 343

R
random order 73, 75
rater 285, 288–289, 291
rating 269, 284–285, 288
reaction time 73, 175
reading process(es) 38, 41, 46
reading times 69
recall 39, 66, 107, 113, 252, 263, 273, 276
reception capacity 187, 193, 197–198
redundancy 249, 255–256, 264
reference materials 152, 156–159, 161–162
reframing 221
register 75, 112–113, 118, 152, 157, 233, 243
regression 40, 96, 98, 101, 104, 108
relative clause 258
reliteralization 33–34
repairs 28, 99, 102, 109–110, 112, 115, 117–118
repetitions 99–101
response latencies 67
responses 178
retrospection 1–2, 39, 269, 273–274, 276–277, 302–303, 306, 351
retrospective data 39, 113
revision 23, 28, 30–34, 40, 94, 99–100, 106, 117, 152, 168, 323, 340, 351
rich points 317, 322–323, 326–328, 330, 333–335, 339

S
scenes and frames theory 221
secondary school 149–150, 152–154, 164–165, 358
secondary school curriculum 165
segmentation 40, 47, 122
self-analysis 57–58, 65
self-assessment 155, 159, 164
semantic differences 183
semantic field 152
semantic organization 184
semantic planning 95
sensorimotor experience 171
sentience 71
shadow translation 26
shift 1, 27–28, 31, 94, 99, 109, 117–118, 155, 244
shift(s) of attention 38, 45
short-term memory (STM) 31, 39–40, 132, 135, 182, 205
sight translation 93–96, 98, 100, 102–104, 116–118
similarity 25–28, 31–33, 114, 222, 227, 255–256
simultaneous interpreting 205
situational factors 302–304, 308, 311–313
skilled memory 249
skilled performance 264
slang 156–158, 162, 187, 190
socially dominant language 81
solidarity 231, 233, 238
source domain 173
source-text reading 38, 40–41
speakerphone 231, 234, 236–237, 239
specificity 201, 213
speech disfluency 93, 102
speech performance 100, 107
standard deviation 78–79, 143
standards 13–14, 17–20
statistical analysis 78–79, 103, 259, 350
statistical reliability 183
strategic sub-competence 319, 323, 326
strategy(ies) 24, 26, 31, 33, 94, 108, 112–116, 118, 122–123, 131, 152–154, 157, 165, 219–220, 222, 226, 228–229, 232, 239, 243, 256, 264–265, 269, 271–272, 274, 277, 283–284, 290–292, 295, 299, 301–304, 306–308, 312, 323, 325–326, 328–329, 339
structural parallelism 219, 228
structural priming 70, 81
student translation 67, 75, 81
subject position 71
subordinate clause 259
subordination 257–259, 261–262, 264
subtitling 187–190, 196–199
summated bigram frequency 74
surtitles 189–199
syllabus 150–152, 154, 165
symbolic activity 20
synonymy 152
syntactic choice 82
syntactic complexity 249–251, 253, 257–258, 265–267
syntactic manipulation 95, 97, 104, 116
syntactic planning 95, 109
syntactic processing 264, 278, 295
syntactic structure 263–264, 295
syntactic transfer disfluencies 109

T
tacit knowledge 301–302, 305–306, 308–312, 315
tags 65–66
target audience 152, 190
target complexity 175
target domain 173

target language 25, 28–29, 31, 33, 58, 60–61, 63–64, 66, 68, 70, 75, 81, 110, 153, 207, 214, 249
target metaphoricity 175
target-text monitoring 38, 45
target-text production 41
target-text reading 38, 45
technical knowledge 19
temporal dynamics 39
tense 58, 125, 152, 163, 208
text 249, 252–256, 258–262, 264
text-awareness 164
text density 190–191
theory of conflation 171
think-aloud 1, 28, 38–39, 57–59, 62, 65, 69, 100, 201, 219, 272, 325, 351–352
tightrope hypothesis 201, 204–205
time lag 121–128, 132, 135–145
trained students 72–73, 75, 79–80, 87–88
translatability 152
translation competence 31, 72, 317–318, 320–321, 325–326, 329–330, 334–335, 339, 352
translation corpora 69, 73
translation didactics 82
translationese 69, 158, 162–163
translation interference 69
translation onset time 77, 82, 87
translation pause(s) 98
translation problem(s) 67–69, 71–72, 78, 80, 93–95, 114, 188, 190, 317–319, 320–323, 325–330, 332–335, 337, 339, 343
translation process(es) 1–2, 24, 26, 28, 31, 37–38, 41, 45, 47, 52, 58, 65, 67, 73–74, 79, 81, 93, 100, 107, 121, 123, 128, 134, 144, 149, 155–157, 162, 165–166, 168–169, 171, 219–221, 223–227, 303, 318–320, 322–323, 325–326, 328, 349, 352
translation project 319–320, 322–323, 326
translation quality 34, 82, 94, 100
translation skills 149–150, 154, 159, 168
Translation Skills Program (TSP) 149–152, 154–157, 159–162, 164
translation strategy disfluencies 108, 112
translation task 39–41, 70, 72, 76, 81, 94, 96, 104, 110, 117–118, 129, 131–134, 144, 150, 152, 221, 228–229, 322–323, 326, 329, 358
translation training curriculum 69
translation workshop 153, 159
translator behaviour 72
translator training 17, 152–153, 163
Translog 37–42, 44, 50–53, 98, 106, 127, 352
TRAP project 39
trials 178
turntaking 15
two-by-two factorial design 73

U
uncertainty 53, 57, 61, 66, 95
unfilled pauses 98–105
unique items 23, 29
universal(s) 23, 29, 30
untrained students 72–73, 75, 80, 87–88
use of instrumental resources 320, 323, 340

V
verbalization 94
visual interference 93, 103–105, 116, 118
visual word recognition 69, 74, 83, 85
voids 152, 157, 160, 162–163
volition 71

W
wiring of neurons 174
within-subject 78, 81
word frequency 74
word length 74
wordplay translation 219
working memory 182, 249, 251–252, 265
writing skills 152
written translation 38, 45–46, 93–96, 100, 104, 106, 114, 121, 129, 149–150, 164, 318