THE PATTERNS OF INTERACTION BETWEEN PROFESSIONAL TRANSLATORS AND ONLINE RESOURCES

Thesis submitted for the Degree of Doctor of Philosophy

JOANNA GOUGH

December 2016

©Joanna Gough
DECLARATION OF ORIGINALITY

I declare that this thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text or bibliography. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.
ABSTRACT

With the rapid growth of the Internet and the recent developments in translation technology, the way translators carry out their translation-oriented research has changed dramatically. Resources used by translators to conduct such research have diversified and largely moved from paper to online. However, whilst the number and the variety of online resources available to translators is growing exponentially, little is known about the interactions between translators and these resources. The present research empirically examines the use of online resources by professional translators during their translation-oriented research activities and it does so from an information behaviour perspective. As a first study of its kind, it focusses on freelance professional translators working at their normal place of work. Specifically, this work addresses the questions of the nature and quantity of resources used by translators as well as the time they spend on research activities. Furthermore, it examines the individual differences between the participants during the research activities. These differences are studied by considering the types of resources used and the ways they are accessed, and by investigating the many volume- and time-related aspects of each translator’s research activities. The main contribution of this study lies in the identification of patterns and their systematisation through a multidimensional analysis, culminating in the formulation of two taxonomies - the Resource Type User Taxonomy (RTUT) and Taxonomy of Translator Research Styles (TTRS). It is argued that whilst RTUT may largely depend on technology developments, TTRS reflects the more innate traits of translators’ information behaviour. By employing a two-stage, multi-method approach (Global Survey, N=540 and Main Study N=16), and by conducting it remotely, through the Internet, the present study represents a quasi-naturalistic research design which aims to observe translation processes as they happen in translators’ natural working environments. This methodology in itself constitutes a contribution to translation process studies.
ACKNOWLEDGEMENTS

I thought this moment would never come... Now that it has, all I can see is not the hours, days, months and years spent in solitude, but the many people who made this moment come true...

First of all, I would like to express my deepest gratitude to my supervisors, Professor Margaret Rogers and Professor Sabine Braun, for the supreme mentoring and invaluable support they have given me throughout this lengthy, but thoroughly satisfying research and writing process. This accomplishment was possible only thanks to their consistent encouragement, goodwill and understanding of my working patterns and the many problems I was faced with during this process.

Secondly, I’d like to thank all the translators who participated in my study, especially those who devoted a considerable amount of time and effort to provide me with data for the Main Study. Your goodwill is highly appreciated and will never be forgotten!

My gratitude is extended to my external examiners, Professor Maureen Ehrensberger-Dow and Professor Gloria Pastor Corpas for their invaluable feedback and for making my Viva such an unforgettable experience which I can only sum up as extremely challenging but utterly satisfying! Also, a big thank you to Dr Ana Frankenberg-Garcia for being my internal examiner and to the lovely Catharine for her relentless proofreading work!

I would also like to thank my children Max and Kayah who have bravely endured all those years of mummy being locked in her study and stressing out about missed deadlines. ‘THE PhD’ will be a banned word in the house from now on. Also, to my husband, Shaun who, I’m sure, will be amongst the ones most pleased with this submission – thank you for your continuous support in my never-ending education!

Last, but not least - my mum, Krystyna, who has been a tremendous help on the domestic front, my Guardian Angel and a constant reminder of that what really matters in life.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................................................... 8
LIST OF TABLES ............................................................................................................................................. 11
1. Introduction .................................................................................................................................................. 14
   1.1. Rationale ................................................................................................................................................. 14
   1.2. Concepts and definitions .......................................................................................................................... 16
   1.3. Overview of Chapters .............................................................................................................................. 19
2. Review of relevant research topics in translation process research (TPR) and related areas .................. 21
   2.1. From black box to tool box (and back again?) ...................................................................................... 21
   2.2. From dictionaries to other types of resources ...................................................................................... 25
   2.3. A growing interest in translators’ online information behaviour ...................................................... 29
   2.4. From laboratory to workplace and novice to expert .......................................................................... 37
   2.5. From academia to industry ................................................................................................................... 40
   2.6. Other relevant research areas ............................................................................................................... 41
       Human translation-related styles .............................................................................................................. 42
       Continuous professional development (CPD) .......................................................................................... 46
       Interdisciplinarity .................................................................................................................................... 46
   2.7. Research questions .............................................................................................................................. 47
3. Methodology ................................................................................................................................................ 49
   3.1. Research paradigm ............................................................................................................................... 49
   3.2. Methods and instruments of data collection and analysis ................................................................ 51
      3.2.1. The evolution of methods of data collection in TPR .................................................................... 51
      3.2.2. Methods of data collection used in the present study ................................................................ 53
      3.2.3. Methods of data analysis used in the present study .................................................................... 60
      3.2.4. Instruments of data collection used in the present study ............................................................ 67
      3.2.5. Instruments of data analysis used in the present study ............................................................... 68
      3.2.6. Materials used in the present study .............................................................................................. 69
   3.3. Research design .................................................................................................................................... 71
      3.3.1. The Global Survey ......................................................................................................................... 72
      3.3.2. The Self-study .............................................................................................................................. 73
      3.3.3. The Pilot Study ............................................................................................................................ 74
      3.3.4. The Main Study ........................................................................................................................... 77
7.1. Categories feeding into the Typology of Translator Research Styles ..............................................194
  7.1.1. Primary categories ..................................................................................................................196
  7.1.2. Secondary categories (attributes) .........................................................................................213
7.2. The Typology of Translation Research Styles (TTRS) .................................................................228
  7.2.1. The Explorer ..........................................................................................................................229
  7.2.2. The Prolific Translator ..........................................................................................................230
  7.2.3. The Methodical Translator .....................................................................................................232
  7.2.4. The Economical Translator ...................................................................................................233
  7.2.5. The Understated Translator ..................................................................................................234
  7.2.6. TTRS Summary ....................................................................................................................235
7.3. The influence of the demographic and professional status variables on RTUT and TTRS ............236
  Length of experience .......................................................................................................................237
  Age .....................................................................................................................................................238
  Gender .............................................................................................................................................240
  Familiarity with the domain .............................................................................................................241
8. Summary of findings and discussion ...............................................................................................243
  8.1. Time spent on research activities ...............................................................................................243
  8.2. The use of resources ..................................................................................................................244
  8.3. Research distribution ..................................................................................................................247
  8.4. Variations in resource behaviour ...............................................................................................249
  8.5. Variations in research behaviour ..............................................................................................253
  8.6. Externalisation of the translation activities .................................................................................256
  8.7. Practical implications of the findings ...........................................................................................257
    Technology development ..................................................................................................................257
    Translator training and teaching .....................................................................................................258
    Research ..........................................................................................................................................259
REFERENCES .................................................................................................................................267
APPENDICES .................................................................................................................................283
LIST OF FIGURES

Figure 1.1 THE PERCEIVED HUMAN TRANSLATION PROCESS (KantaxMT, 2010: online).................................16
Figure 2.1 A VIEW OF THE TRANSLATOR’S ‘WORKSPACE’ PRE- AND POST- TRANSLATION TECHNOLOGY.....23
Figure 2.2 A NESTED VIEW OF INFORMATION BEHAVIOUR (Wilson, 1999).............................................30
Figure 2.3 DIMENSION OF EXPLORATION AND SPECIFICITY RELATED TO INFORMATION SEEKING STYLE,
STUDY APPROACH AND PERSONALITY - Figure adapted from Heinström (2006a:1449)..........................32
Figure 3.1 THE STUDY COMPONENTS AND ANALYTICAL APPROACH................................................49
Figure 3.2 METHODS OF DATA COLLECTION ADAPTED FROM Krings (2005)........................................52
Figure 3.3 METHODS OF DATA COLLECTION AS EMBEDDED IN THE OVERALL FLOW OF DATA .............54
Figure 3.4 METHODS OF DATA ANALYSIS AS EMBEDDED IN THE OVERALL FLOW OF DATA ..................61
Figure 3.5 SCREENSHOTS OF TEXT AND RESOURCE WINDOWS ..........................................................62
Figure 3.6 SIMULTANEOUS DISPLAY OF TEXT AND RESOURCE WINDOW .............................................63
Figure 3.7 THE STRUCTURE OF A RESEARCH ACTION ...........................................................................63
Figure 3.8 ‘SNIPPETS’ IN GOOGLE SEARCH ENGINE RESULTS ............................................................64
Figure 3.9 INSTRUMENTS OF DATA COLLECTION AS EMBEDDED IN THE OVERALL FLOW OF DATA ..........67
Figure 3.10 INSTRUMENTS OF DATA COLLECTION AS EMBEDDED IN THE OVERALL FLOW OF DATA ........68
Figure 3.11 RESEARCH DESIGN FLOWCHART .........................................................................................71
Figure 3.12 GLOBAL SURVEY DESIGN .................................................................................................72
Figure 3.13 SELF-STUDY DESIGN ........................................................................................................73
Figure 3.14 PILOT STUDY DESIGN .......................................................................................................74
Figure 3.15 MAIN STUDY DESIGN ....................................................................................................78
Figure 3.16 THE STRUCTURE OF THE PRESENT STUDY - FROM RESEARCH DESIGN TO HYPOTHESIS ....79
Figure 4.1 A SNAPSHOT OF GOOGLE SEARCH SHOWING DIFFERENT REPRESENTATIONS OF TRANSLATION
TECHNOLOGY ........................................................................................................................................85
Figure 4.2 TRANSLATION TECHNOLOGIES - TOOLS AND RESOURCES ................................................87
Figure 4.3 DISTINCTION BETWEEN TOOLS AND RESOURCES EMPHASISING THE BLURRED BOUNDARIES
BETWEEN THEM ..................................................................................................................................87
Figure 4.4 BABELNET – A MULTILINGUAL ENCYCLOPAEDIC DICTIONARY AND A SEMANTIC NETWORK
INCORPORATING 6 DIFFERENT RESOURCES ........................................................................................88
Figure 4.5 THE DIFFERENCES BETWEEN TOOLS AND RESOURCES .......................................................89
Figure 4.6 ONLINE RESOURCES IN A WIDER CONTEXT OF DIGITAL TECHNOLOGIES .............................90
Figure 4.7 BABELNET - A DICTIONARY, ENCYCLOPAEDIA AND ONTOLOGY .........................................93
Figure 4.8 TYPES OF RESOURCES FOR TRANSLATORS ........................................................................94
Figure 4.9 DONUT GLOSSARY .............................................................................................................96
Figure 4.10 WIKIPEDIA GLOSSARIES ................................................................................................97
Figure 4.11 SKETCH ENGINE .................................................................................................................100
Figure 4.12 SKELL – A SCALED-DOWN VERSION OF SKETCH ENGINE ....................................................101
Figure 4.13 SNAPPY WORDS – A VISUAL THESAURUS ..........................................................................102
Figure 4.14 THE CONCORDANCING FUNCTION IN SKETCH ENGINE .......................................................102
Figure 4.15 WOLFRAM ALPHA – EXAMPLES OF QUERIES IN THE WORDS & LANGUAGES CATEGORY ....104
Figure 4.16 WOLFRAM ALPHA – CATEGORIES IN WHICH QUERIES CAN BE SUBMITTED FOR COMPUTATION
BASED ON BUILT-IN DATA, ALGORITHMS AND METHODS ..................................................................104
Figure 4.17 AN EXAMPLE OF A KEYWORD SEARCH ENGINE ................................................................109
Figure 4.18 AN EXAMPLE OF A DIRECTORY SEARCH ENGINE .......................................................... 109
Figure 4.19 AN EXAMPLE OF A HYBRID SEARCH ENGINE .......................................................... 110
Figure 4.20 THE ‘BITCOIN’ WEBSITE CONTAINING A GLOSSARY OF RELEVANT TERMS .................. 111
Figure 4.21 TYPOLOGY OF TRANSLATION AIDS. ...................................................................... 115
Figure 4.22 BASIC DIVISION OF TRANSLATION AIDS INTO DICTIONARIES AND OTHER REFERENCE AIDS .... 115
Figure 4.23 MACMILLAN DICTIONARY’S OPEN DICTIONARY SECTION ....................................... 116
Figure 5.1 PROFESSION OF PARTICIPANTS FROM THE GLOBAL SURVEY AND THE MAIN STUDY .......... 125
Figure 5.2 GEOGRAPHICAL LOCATION OF PARTICIPANTS FROM THE GLOBAL SURVEY AND THE MAIN STUDY ................................................................. 126
Figure 5.3 AGE DISTRIBUTION IN THE GLOBAL SURVEY AND THE MAIN STUDY ............................. 126
Figure 5.4 GENDER DISTRIBUTION - COMPARISON OF DATA FROM GLOBAL SURVEY AND MAIN STUDY WITH ITI STATISTICS.................................................................................. 127
Figure 5.5 AGE AND GENDER DISTRIBUTION IN THE GLOBAL SURVEY AND MAIN STUDY .............. 127
Figure 5.6 SELF-DECLARED SPECIALISATION IN THE GLOBAL SURVEY AND MAIN STUDY .................. 129
Figure 5.7 PROPORTION OF MAIN STUDY PARTICIPANTS SPECIALISING IN BUSINESS/ECONOMIC/FINANCIAL DOMAIN ........................................................................................................... 130
Figure 5.8 PROPORTION OF MAIN STUDY PARTICIPANTS COMFORTABLE WITH THE SOURCE TEXT WHICH BELONGED TO BUSINESS/ECONOMIC/FINANCIAL DOMAIN .............................................................. 130
Figure 5.9 SELF-DECLARED AND OBSERVED TIME SPENT ON RESEARCH GLOBAL SURVEY AND MAIN STUDY DATA .............................................................................................................. 133
Figure 5.10 OBSERVED PERCENTAGE OF TIME SPENT ON RESEARCH IN THE MAIN STUDY ................ 134
Figure 5.11 THE RELATIONSHIP BETWEEN THE USE OF TRANSLATION MEMORY SYSTEMS, TERMINOLOGY MANAGEMENT SYSTEMS AND NON-CUSTOMISED MT SYSTEMS AND THE TIME SPENT ON RESEARCH ACTIVITIES – SELF-DECLARED DATA FROM THE GLOBAL SURVEY ........................................... 135
Figure 5.12 QUESTION 8 OF THE GLOBAL SURVEY (THE USE OF ONLINE, PAPER AND OTHER RESOURCES) 136
Figure 5.13 EXAMPLES OF GLOBAL SURVEY RESPONDENTS’ INDICATION OF TIME SPENT ON ONLINE, PAPER & OTHER RESOURCES ........................................................................................................ 136
Figure 5.14 SELF-DECLARED PERCENTAGE OF USE OF VARIOUS TYPES OF RESOURCES ACCORDING TO MEDIUM (ONLINE, PAPER & OTHER) - GLOBAL SURVEY DATA .......................................................... 137
Figure 5.15 SELF-DECLARED USE OF TYPES OF RESOURCES IN THE GLOBAL SURVEY AND IN THE MAIN STUDY .............................................................................................................. 138
Figure 5.16 THE USE OF DICTIONARIES IN THE GLOBAL SURVEY AND IN THE MAIN STUDY ................ 139
Figure 5.17 SELF-DECLARED USE OF MONOLINGUAL AND BILINGUAL DICTIONARIES IN THE GLOBAL SURVEY .................................................................................................................. 139
Figure 5.18 SELF-DECLARED USE OF ONLINE GLOSSARIES, TERMBASES, CORPORA & TM IN THE GLOBAL SURVEY AND IN THE MAIN STUDY .................................................................................. 140
Figure 5.19 NUMBER OF TYPES OF RESOURCES USED BY INDIVIDUAL PARTICIPANTS IN THE MAIN STUDY 143
Figure 5.20 THE MOST USED SPECIFIC RESOURCES SELF-DECLARED IN GLOBAL SURVEY .................... 145
Figure 5.21 THE MOST USED SPECIFIC RESOURCES SELF-DECLARED IN MAIN STUDY ...................... 145
Figure 5.22 THE LEAST USED SPECIFIC RESOURCES SELF-DECLARED IN GLOBAL SURVEY ................. 146
Figure 5.23 THE LEAST USED SPECIFIC RESOURCES SELF-DECLARED IN THE MAIN STUDY .................. 146
Figure 5.24 FINDING A SOLUTION TO THE TRANSLATION OF THE TERM FIAT MONEY IN ONE STEP USING 2LINGUAL .................................................................................................................. 147
Figure 5.25 CONFIRMING A SOLUTION TO THE TRANSLATION OF THE TERM FIAT MONEY IN ONE STEP USING 2LINGUAL .............................................................................................................. 148
LIST OF TABLES

Table 4.1 DEDUCTIVE CLASSIFICATION OF TRANSLATION RESOURCES ACCORDING TO ‘CHOICE’: A SUGGESTED CATEGORISATION ................................................................. 116
Table 4.2 LIST OF RESOURCE TYPES AND THEIR USES. Source: Désilets et al., 2009 ................................. 117
Table 4.3 VARIOUS POSSIBLE CLASSIFICATIONS OF TRANSLATION RESOURCES ................................................. 118
Table 5.1 AGE DISTRIBUTION IN THE GLOBAL SURVEY AND THE MAIN STUDY ............................................. 127
Table 5.2 LENGTH OF EXPERIENCE DISTRIBUTION IN THE GLOBAL SURVEY AND MAIN STUDY .............. 128
Table 5.3 SPECIALISATIONS OF THE GLOBAL SURVEY PARTICIPANTS ................................................................. 128
Table 5.4 SPECIALISATIONS OF THE MAIN STUDY PARTICIPANTS .................................................................. 129
Table 5.5 THE FREQUENCY OF USE OF ALL TYPES OF RESOURCES IN THE MAIN STUDY ................................. 143
Table 5.6 BREAKDOWN OF STEPS TAKEN BY INDIVIDUAL TRANSLATORS TO REACH THEIR CONCLUSION REGARDING THE TRANSLATION OF THE TERM FIAT MONEY ...................................................... 147
Table 5.7 LANGUAGES REPRESENTED IN THE MAIN STUDY ........................................................................... 149
Table 5.8 FREQUENCY OF USE OF THE MOST USED COMMON RESOURCES ...................................................... 150
Table 5.9 EXAMPLE SHOWING THE DIFFERENCE BETWEEN THE NUMBER OF STEPS AND THE NUMBER OF SPECIFIC RESOURCES .................................................................................. 154
Table 5.10 MOST FREQUENTLY USED COMMON RESOURCES OBSERVED IN THE MAIN STUDY .................. 158
Table 5.11 THE LENGTH OF RESEARCH STEP ..................................................................................................... 164
Table 6.1 CATEGORIES AND SUB-CATEGORIES OF THE RESOURCE TYPE USER TYPOLOGY (RTUT) ............ 174
Table 6.2 CATEGORIES OF RESOURCES USED IN THE PRESENT STUDY ACCORDING TO THEIR NATURE ...... 175
Table 6.3 THE PROCESS OF ESTABLISHING RESOURCE DOMINANCE .............................................................. 176
Table 6.4 DISTRIBUTION OF TERMO-LEXICOGRAPHIC AND TEXT-BASED RESOURCES ........................ 177
Table 6.5 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NATURE OF RESOURCES USED .......................................................................................................................... 177
Table 6.6 SUMMARY OF CLASSIFICATION OF TRANSLATORS ACCORDING TO THE TYPE OF RESOURCES THEY USE .................................................................................................................. 178
Table 6.7 EXAMPLE OF THE ST-ORIENTED AND TT-ORIENTED RESEARCH DIRECTIONS BEING ASSOCIATED EXCLUSIVELY WITH THE TYPE OF RESOURCE ............................................................................. 180
Table 6.8 EXAMPLE OF HOW RESEARCH DIRECTION CODES WERE ASSIGNED IN RESEARCH EPISODES .... 181
Table 6.9 DISTRIBUTION OF ST-ORIENTED AND TT-ORIENTED SEARCHES ................................................... 182
Table 6.10 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE RESEARCH DIRECTION ... 182
Table 6.11 EXPERIENCE VS. RESEARCH DIRECTION ............................................................................................. 183
Table 6.12 VARIOUS CLASSIFICATIONS OF TYPES OF SEARCH QUERIES ......................................................... 184
Table 6.13 DISTRIBUTION OF ST-ORIENTED AND TT-ORIENTED SEARCHES ................................................... 189
Table 6.14 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO RESEARCH STRATEGY ......... 189
Table 6.15 COLOUR CODING OF CATEGORIES IN RTUT ....................................................................................... 190
Table 6.16 HEATMAP SHOWING HOW CATEGORIES WERE ASSIGNED TO INDIVIDUAL TRANSLATORS TO FORM RTUT .................................................................................................................. 190
Table 7.1 THE PRIMARY TTRS CATEGORIES AND THEIR SUB-CATEGORIES ...................................................... 194
Table 7.2 THE SECONDARY TTRS CATEGORIES .................................................................................................. 195
Table 7.3 OVERVIEW OF PRIMARY CATEGORIES AND THEIR COMPONENTS .................................................. 196
Table 7.4 COLOUR CODING AND CHARACTERISING BEHAVIOURS OF THE CATEGORIES FOR THE TYPOLOGY OF TRANSLATOR RESEARCH STYLES .......................................................................................... 197
Table 7.5 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NUMBER OF RESEARCH UNITS RESEARCHED IN THE TASK ................................................................. 198
Table 7.6 RU VOLUME CLASSIFICATION ................................................................. 198
Table 7.7 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NUMBER OF SPECIFIC RESOURCES USED IN THE TRANSLATION TASK ................................................. 199
Table 7.8 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE FREQUENCY OF USE OF SPECIFIC RESOURCES IN THE TRANSLATION TASK ............................................. 200
Table 7.9 RESEARCH VOLUME CLASSIFICATION .................................................... 201
Table 7.10 THE NUMBER OF STEPS TAKEN BY INDIVIDUAL TRANSLATORS TO SATISFY RESEARCH NEEDS RELATED TO THE TERM ‘FIAT MONEY’ ................................................................. 201
Table 7.11 COMPARISON OF RESEARCH STEPS OF T13 AND T16 ................................ 202
Table 7.12 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO RESEARCH INTENSITY ............... 203
Table 7.13 RESEARCH INTENSITY CLASSIFICATION ............................................. 203
Table 7.14 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NUMBER OF TYPES OF RESOURCES USED IN THE TRANSLATION TASK ............................................. 205
Table 7.15 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NUMBER OF SPECIFIC RESOURCES USED PER RESEARCH UNIT ......................................................... 206
Table 7.16 CLASSIFICATION OF TRANSLATORS ACCORDING TO RESEARCH DIVERSITY ....................... 206
Table 7.17 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO TIME SPENT ON RESEARCH (MEASURED IN REAL TIME) ................................................................. 207
Table 7.18 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO TIME SPENT ON RESEARCH (MEASURED IN PROPORTION TO TRANSLATION TIME) ............................................. 208
Table 7.19 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE AVERAGE DURATION OF RESEARCH STEPS .................................................................................... 210
Table 7.20 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE AVERAGE DURATION OF RESEARCH EPISODES .................................................................................... 211
Table 7.21 RESEARCH PACE CLASSIFICATION ....................................................... 211
Table 7.22 THE PRIMARY CATEGORIES REPRESENTING THE VOLUME AXIS ....................... 211
Table 7.23 THE PRIMARY CATEGORIES REPRESENTING THE TIME AXIS ............................. 211
Table 7.24 SCORE CALCULATION FOR THE VOLUME AND TIME AXIS OF THE TTRS ....................... 212
Table 7.25 WEIGHT SYSTEM FOR ASSIGNING SECONDARY CATEGORIES (ATTRIBUTES) TO TRANSLATORS’ PROFILES .................................................................................................. 214
Table 7.26 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO DEEP SEARCHES ................ 214
Table 7.27 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO SHALLOW SEARCHES ............. 215
Table 7.28 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO MEANDERING RESEARCH PATHS ............................................................................................................. 216
Table 7.29 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO RESEARCH SEQUENCE ............. 217
Table 7.30 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO PLANNING ATTRIBUTE ............ 218
Table 7.31 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO SQUIRRELLING BEHAVIOUR ATTRIBUTE ....................................................................................................... 219
Table 7.32 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO SNIPPET VIEWING ATTRIBUTE ............................................................................................................. 220
Table 7.33 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO ADVANCED SEARCH QUERIES .................................................................................................................. 221
Table 7.34 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO STRATEGIC BEHAVIOUR ......... 222
Table 7.35 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO DRIVE FOR PERFECTION ........224
Table 7.36 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO AFFECTIVE ASSESSMENT ..........225
Table 7.37 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO HIGH INFORMATION RETENTION .......................................................... 225
Table 7.38 INDIVIDUAL TRANSLATORS' ATTITUDE TOWARDS TECHNOLOGY ........................................... 226
Table 7.39 THE INFLUENCE OF EXPERIENCE ON THE RTUT and TTRS COMPONENTS ........................................... 237
Table 7.40 THE INFLUENCE OF EXPERIENCE ON RTUT AND TTRS ................................................................. 238
Table 7.41 THE INFLUENCE OF AGE ON THE RTUT AND TTRS COMPONENTS .................................................. 239
Table 7.42 THE INFLUENCE OF AGE ON RTUT AND TTRS ................................................................................ 239
Table 7.43 THE INFLUENCE OF GENDER ON THE RTUT AND TTRS COMPONENTS ............................................. 240
Table 7.44 THE INFLUENCE OF GENDER ON RTUT AND TTRS ........................................................................ 240
Table 7.45 THE INFLUENCE OF DOMAIN FAMILIARITY ON THE RTUT AND TTRS COMPONENTS .................. 241
Table 7.46 THE INFLUENCE OF DOMAIN FAMILIARITY ON RTUT AND TTRS .................................................. 241
1. Introduction

To understand how professional translators interact with online resources, we need to acknowledge that within the span of twenty years or so, translation in professional settings has transformed beyond recognition. The vast majority of online resources have moved from research and development labs to the market and translation as a global industry has become very much dependent upon translation technology, of which online resources form a vital part. In an account from a ProZ\textsuperscript{1} member, a picture of what it was like to carry out research in the pre-Internet days becomes very vivid:

I had a ton of paper dictionaries (over 300) and I would sit at a table with all of the appropriate dictionaries open (business, legal, etc.). If the word or phrase wasn’t in any of your paper dictionaries, you were in big trouble. You had to spend hours and hours looking things up. [...] I also had a book full of paper glossaries, clippings of word lists from magazines, the ATA Chronicle, terms on index cards. [...] I would also have to make frequent trips to the university library and make telephone calls to experts in order to ask questions.  

(ProZ, 2015b)

Nowadays the list of resources available online in a large number of languages will most likely be longer than the present thesis. The Internet itself has become the biggest resource of all; a goldmine of knowledge and information contained in its web pages and the biggest source of tools and specialised resources for translators to help them access this knowledge at the click of a mouse. As Enríquez Raído observes, “[t]he impact and penetration of the Internet have indeed dramatically transformed the way translators carry out their documentary research and address their information needs” (2013:2). The present thesis is concerned with this transformation and seeks to find out how freelance professional translators are engaging with the online technological offerings of the present age, the Information Age, to satisfy their research needs during the translation process.

1.1. Rationale

The various technologies supporting translators in their work are rapidly evolving and the adoption of these technologies by translators is on the increase (Fulford & Zafra, 2004; O’Hagan, 2013:504). Therefore, there is a need for research in translation studies to examine the use of technology by translators and the impact it has on their work. The fast-paced development of technology happening in so many diverse and yet interlinked areas necessitates increased efforts from the research communities to capture these developments and their impact on translation process and product.

However, in order for this technology to address the needs of translators, the goals of such increased research efforts have to go beyond the practical aims of surveys aiming to find out which particular technologies are being used and beyond research designs that aim at improving the translation tools or resources, increasing the leveraging of data and productivity. Such research needs to consider the deeper implications of how technology affects the work of translators and to reflect upon the psychological, social or ethical aspects of these implications.

\footnote{1 http://www.proz.com/}
Furthermore, if technology is to serve human beings, research needs to go beyond the productivity-driven objectives and consider the more individual aspects of human-computer interaction, embracing diversity and individuality in this interaction in order to gain “better knowledge about the structure of individual translation processes and personal translation styles” (Hansen, 2013:88). As O’Brien points out, the relationship between personality and the translation process has received little attention and “intriguing questions that could be posed […] have not yet been investigated” (2013:8). Deeper and broader research is therefore vital to ensure that translation technology does not evolve independently of the translator, but that research findings are fed into its development in order to serve the needs of the translator, not the other way round. Ehrensberger-Dow and Massey have already noted that “translators might be adapting their performance to software rather than the converse” (2014:80).

And finally, research into the use of technology in the translation process is vital to the understanding of the various aspects of the relationship between translators and technology in order to provide guidelines for teaching and training of the new generation of translation professionals entering an industry that has been recently so dramatically transformed. Research is needed especially in the areas of translation competence that are most affected by a technological change such as information competence (EMT Expert Group, 2009) technological competence (EMT Expert Group, 2009) or instrumental competence (PACTE, 2009). The present research is an attempt to address all the above needs by focussing on the use of translation technology by professional translators to carry out their research activities during the translation process.

The human translation process comprises two main observable activities: interactions with the texts (source and target) and interactions with resources. Although research into the translation process dates back a few decades, these two activities have not been explicitly separated and researched, i.e. the use of external resources was hitherto examined as a translation activity, not an information seeking activity. As Enríquez Raído observes, the general trend that can be seen in the studies carried out in the 1980s and 1990s was “to examine translation strategies that draw on both internal and external resources” (Enríquez Raído, 2011:71-72) and so they examine both internal and external processes as one problem-solving activity. In some cases (e.g. Lorscher, 1986), external resources were even disallowed during the experiments, much to the disapproval of Bernardini (2001:260) and Kiraly (1995:44-45), who remarks that “there is much to be learnt about the research aspect of translation”. Pinto and Sales (2008:433) also emphasise the fundamental role of research in the translation process.

The industry’s focus has been on a related issue, i.e. developing technology that aids translation activities, whilst the role of research activities and how these can be supported has been largely ignored (see section 2.4.). Risku aptly portrays translators who “in their search for coherent meanings, […] assume the roles of the researchers and archaeologists […] to produce communicatively or pragmatically appropriate texts”. This comparison provides a lucid image of what translators’ activities might look like during their interaction with the resources that enable the production of a target text (2010:97). And yet, as seen in Figure 1.1, the complexity of the translation process is not always recognised, especially outside academia, where the translation process can be portrayed in simplistic terms as a linear process from the source text to the target text, with the translation process occurring solely in the translator’s head (KantanMT, 2010).
Such a portrayal ignores the complex research activities that form an important part of the translation process. In translation studies too, while process-oriented research has been focusing on internal translation processes (taking place in the translators’ heads while they are translating), not enough attention has been paid to how translators access and utilise the external resources available to them, especially those available online. This research gap is further exacerbated by the fact that whilst the amount of resources available to translators is growing exponentially, still relatively little is known about the interaction between translators and these resources. Therefore, it is surprising that such an integral part of the translation process has received such little attention from research communities.

The present research aims to fill this gap by exploring the translation process as a complex and intertwined series of interactions with texts (source and target) and interactions with external resources (reference material), but from the perspective of the resources. Although, from the cognitive point of view, both of these types of interactions constitute a different activity and involve different mental processes, they are intertwined and integrated into the state of a *flow experience*[^2], in which translators remain in control of the process and product and are able to monitor themselves (Hansen, 2003:26).

The present research, whilst addressing the practical issues such as time spent on research activities and the types of resources used in these activities, also attempts to recognise the implications of how humans interact with computers on an individual basis, encompassing individual differences and research styles.

### 1.2. Concepts and definitions

The aim of this section is to provide the definitions of key concepts that are referred to in this thesis in order to clarify the scope and perspective of the study, as many of these concepts are relatively unstable and open to interpretation.

[^2]: This complexity is further compounded in the post-editing process, where an additional phase, “suggested translation evaluation” (Green et al., 2013) is now considered to constitute part of the process.

[^3]: The concept of a *flow experience* was proposed by a Hungarian-American psychologist Mihaly Csikszentmihalyi and was incorporated by Hansen (2003: 26) in the context of the translation process.
Translation process
The definition of the translation process is adapted from Hansen (2003:26) who refers to it as:

   everything that happens from the moment the translator starts working on the source text until he finishes the target text. It is all encompassing, from every pencil movement and keystroke, to dictionary use, the use of the internet and the entire thought process that is involved in solving a problem or making a correction - in short everything a translator must do to transform the source text to the target text.

(Hansen, 2003: 26)

Therefore, it could be said that the translation process encompasses three types of interactions:
- interactions with the texts (source and target),
- interactions with external resources (translation aids and reference materials) and
- interactions with internal resources, i.e. the translator’s own memory. However, since no cognitive insights such as examining thought processes, or decision making are being sought in this study, the interactions with internal resources will not be examined here; the translation process will only be discussed in terms of what can be observed, here using screen recordings, i.e. the translator’s interactions with the texts and the interactions with external online resources. Therefore, henceforth, when referring to the translation process, it will be assumed that it is composed of these two types of interactions:
  - interactions with source and target texts
  - interactions with external resources.

Translation-oriented research
Translation-oriented research refers to the second type of activity as defined above, i.e. interactions with resources. It is understood to be an integral part of the translation process but requires skills which are different from translation skills, i.e. information-seeking skills.

Translation task
The translation task is understood to consist of a succession of what I am calling translation episodes (periods of interactions with the source or the target text during text production) and research episodes (periods of interactions with online resources). Translation episode and research episode are defined more fully in section 3.2.3.

Professional translator
The term professional translator is used by various researchers but with different definitions. Saldanha and O’Brien lament that this makes cross-study comparisons difficult and “the state of affairs seriously hampers the building, testing and validating of models and theories in the discipline” (2014:114). This becomes even more difficult with the newly emerging models of translation such as community translation or even machine translation with post-editing. Added to this mix of difficulties is Jääskeläinen’s (2010) argument that not all professional translators produce expert translations, and therefore research studies need to be reconsidered in the light of these two notions (professionalism and expertise). Since the present study does not assess the quality of any translations produced by its subjects, no link can be made between expertise and professionalism.
Therefore, for the purpose of this study, professional translators are defined as those who have worked as translators as their main occupation for several years, without taking into account any translation-related education or qualifications. I refer to the participants of the observational component (the Main Study) as experienced professional translators as they were selected on the basis of at least five years of working as translators as their main occupation (although some of them worked part-time at the time of contributing to the present study or have worked part-time, but for over 5 years).

**Translation professional**

*Translation professional* in this study refers to someone who works in the translation industry, but not necessarily or solely as a translator. This term has been used to refer to the participants in the preliminary survey component of the present study which includes translators, interpreters, project managers, terminologists, educators/trainers, localizers, translation business owners and language technology specialists (see section 5.1.2.).

**Translation problem**

In this work *translation problem* refers to any difficulty encountered in the source text that causes the translator to consider options as to the rendering of a particular section of the source text in the target text using internal or external resources. In this study, the notion of translation problem is closely tied to the notion of *research unit* (the particular section of the source text) and to the *research need* (the type of information the translator needs to overcome the problem related to this particular section of the source text) - see section 3.2.3 for more detailed definitions of ‘research unit’ and ‘research need’.

The notions of a translation problem and research need are closely related, but they differ in what follows their identification in the source text. According to Nord (2001:64) “translation problems will always remain problems, even when a translator has learnt how to deal with them rapidly and effectively”. This ‘dealing’ with the problem could include automating of the process or the use of internal resources. However, many translation problems require further action using external resources. In this study, these translation problems that require further investigation using external resources become *research needs* (see section 3.2.3. for a definition of research need).

**Resources**

*Resources* are understood to be any aids that translators use in response to translation problems as defined above. Translators can use *internal resources* (i.e. their human memory) to make decisions unconsciously and automatically or with some deliberation or they can use *external resources* (i.e. any available paper, electronic or online resources, as well as consulting other humans) when internal resources do not yield a satisfying option to solve their problem.
1.3. Overview of Chapters

The present thesis is composed of eight Chapters. Chapter 1 has presented the rationale for the present research and introduced some concepts and definitions to which references are made throughout this work.

Chapter 2 presents a review of literature that is of relevance to the present work. It sets the present study within a wider research context in which research gaps are identified thus leading to the formulation of the research questions. Various research trajectories, triggered by technological developments and the expansion of the Internet, are discussed with a view to providing a justification for the choice of subject for the present study and the attendant methodology. Literature from translation process studies as well as from information behaviour studies is examined in the light of these research trajectories, which include: the shift from the sole investigation of mental processes within translation process, the growing interest in studying the use of translation technology and translators’ online behaviour, the shift from carrying out studies in laboratory conditions to studying translators in their natural working environments and the growing interest in research with an industry focus. Other areas of relevance to the present study discussed in Chapter 2 include research into human translation styles, continuous professional development and interdisciplinarity. The review shows the need for a study of current translation-oriented research practice which assumes a qualitative change in the way translators work.

Based on the choices made with regard to the research questions of the study that emerged from the literature review, Chapter 3 describes the methodology that was developed to investigate these research questions in the most unobtrusive way in order to achieve the highest possible ecological validity. The Chapter presents the various methods and instruments of data collection and analysis, followed by a description of study components, i.e. the Global Survey and the Main Study which were preceded by a pilot phase. Chapter 3 concludes with a discussion on methodological challenges encountered during the study.

Chapter 4 deals with the technological background for the present study. It argues for a more explicit treatment of online resources as part of translation technology and a more coherent way of treatment of translation technology within a wider context of digital technology. It also provides an overview of the online resources currently available for translators and as well as their various classifications.

Chapter 5 proceeds to analyse the self-declared data from the Global Survey and the Main Study, and the observed data obtained from the Main Study. The Chapter highlights the common tendencies with regard to the use of online resources by translators taking part in the present study and compares tendencies identified in the Global Survey with those identified in the Main Study. After analysing the two data sets, the general patterns of the use of resources by the participants are discussed, including the time spent on research activities, the use of paper vs. online resources, the types of resources used as well as which specific resources are consulted and how frequently. This is followed by an examination of research distribution patterns, including the micro-patterns, i.e. relating to one translation problem, and macro-patterns, relating to the whole translation task.

Chapters 6 and 7 move from the common tendencies to focus on the observed individual differences between the sixteen translators that took part in the Main Study, with the aim of highlighting these differences as well as identifying patterns of behaviour within this sample. Chapter 6 is concerned with what I have called translators’ resource behaviour, i.e. the types of resources they choose to access
(termino-lexicographic vs. text-based), the way they access them (via a keyword in a search engine or via a known resource) and the accompanying research direction (i.e. source text vs. target-text oriented). The analysis of these three aspects of translators’ resource behaviour results in the Resource Type User Taxonomy (RTUT) in which three resource type users were identified: the Dictionary Enthusiast, the Parallel Text Fan and the Mixed User, with an additional category of MT user.

Chapter 7 focuses on another aspect of the individual differences that were observed in the Main Study, i.e. the translators’ research behaviour. The Chapter analyses various volume- and time-related aspects of the participants’ online research behaviour using Primary Categories such as the number of lexical items researched, the number and variety of resources used to research those items, and the time and pace of research. Furthermore, fifteen so-called Secondary Categories of a more qualitative nature such as the depth of searches or the ‘shape’ of the search paths are considered to complement the Primary Categories, concluding with the presentation of the Taxonomy of Translator Research Styles (TTRS), in which five research styles were identified: the Economical, the Methodical, The Understated, the Prolific, and the Explorer. Chapter 7 ends with a brief analysis of the influence of variables such as age, gender, length of experience and familiarity with the domain on the two taxonomies.

Chapter 8 presents a summary of the key findings along with a discussion of their significance and their practical implications for technology development, translator training and further research. It also summarises the contribution of this work to translation studies as well as its limitations.

The main contribution of this thesis will be shown to be the identification of patterns of resource and research behaviour within the group of the sixteen freelance professional translators and the methodology developed to discern, systematise and cluster these patterns which led to the formulation of two taxonomies - one representing the resource style (RTUT) and one representing the research style (TTRS).
2. Review of relevant research topics in translation process research (TPR) and related areas

Over the last couple of decades, translation process research has undergone a substantial transformation. In her article ‘Trajectories of Research in Translation Studies’ published a decade ago, Tymoczko identifies six broad areas of “principle trajectories of research in translation studies that are likely to continue and to be productive in the coming years” (Tymoczko, 2005:1082). One of these ‘trajectories’ is identified as “Responses to Technological Shifts and Globalisation” (Tymoczko, 2005:1088). The present research thematically fits into this trajectory as it examines the use of technology, in particular online resources, by professional translators in the globalised world. It is argued that the professionalisation, technologisation and industrialisation of the translation process together with the growth and expansion of the Internet has had an impact on translation practice and consequently on the way translators carry out their translation-related research. This, in turn, necessitates a response from the research communities to examine this impact and the attendant changes in translation practice.

The following review examines the current research trajectories from which the ideas and motivations for the present study are drawn. By doing so, an attempt is made to position the research questions of the present work within the wider context of the current research landscape and to justify the decisions made with regard to the objectives of the present study and the applied methodology. These decisions will be discussed in the context of the following directions:
- from black box to tool box
- from dictionaries to other types of resources
- from translation behaviour to information behaviour
- from laboratory to workplace and from novice to expert
- from academia to industry.

This will be followed by a brief summary of other selected research areas relevant to the present study such as the investigations into human translation styles, interdisciplinarity in translation studies research and the continuous development of technology-related skills and competencies for translators.

2.1. From black box to tool box (and back again?)

From a diachronic perspective, it could be argued that the technologisation of the translation process has had a profound impact on the research trajectories in translation studies and has provided a strong impetus to the growth in the volume of research in process-oriented studies. With special relevance for the present study, one particular shift, from ‘black box’ to ‘tool box’, can be observed in which the examination of translators’ mental processes during the translation process (defined in section 1.2.) has shifted towards the observation of what is happening ‘externally’ to this process.

The first signs of these changes could be seen in the change of research direction from prescriptive, theoretical and speculative product-oriented research to process-oriented research characterised by descriptive, performance-analytical and empirical inquiry (Lörscher, 1992:146), which shifted the

---

4 The present research was carried out in its entirety through the Internet, with participants located in Argentina, Belgium, Brazil, Hungary, Indonesia, Spain, Switzerland, UK and US.
perspective from the outcome of the translation process to the process itself. However, in the early days of this new research direction the object of investigation, the translation process, was still relatively unaffected by translation technologies. Therefore, the focus was mostly directed at the internal support, i.e. the processes happening in the translator’s ‘black box’, rather than on any external support such as the use of tools and resources supporting this process. With time, and the gradual ‘externalisation’ (Cronin, 2003; Pym, 2011) of certain activities within the translation process through assistive technologies such as online resources, translation memory or machine translation, the distinction between internal and external support became more evident, and a shift towards a new ‘growth area’ within process-oriented studies, i.e. external support, became apparent.

Therefore, it could be argued that the technologisation of the translation process has highlighted the dichotomy between internal and external support (Alves & Liparini Campos, 2005) which is often referred to in process-oriented translation research (e.g. Prassl, 2010; Valli, 2012; Enriquez Raído, 2014), where internal support refers to a person’s existing memory, knowledge and past experiences, while external support relates to the use of translation tools and reference material such as dictionaries, encyclopaedias or other external sources. Other scholars draw a similar distinction, e.g. Schubert comments that internal processes relate to “the mental activity involved in carrying out the translation work with all its steps and decisions [while the] external processes are everything in the translation process which can be observed by another person” (Schubert, 2009:19). Dam-Jensen also draws on this distinction and classifies thoughts and decisions as internal processes and the use of tools or communication with clients as external processes (Dam-Jensen, 2012:150). In the present work, internal support is linked to the concept of the ‘black box’, i.e. to all cognitive processes and the use of the translator’s human memory and past experiences, whilst external support is linked to the concept of the ‘tool box’, i.e. translation technology, consisting of tools and resources.

It could be argued that after the distinction between internal and external support had been drawn in TPR literature, a status quo was established in which translation technologies were recognised as being important, and in fact, inseparable from the translation process, thus attracting more interest from research communities. It could also be said that, implicitly, this shift signalled a change in the perception of the translation process as something that happens solely in a translator’s head to a process where the translator’s (human) memory is increasingly extended and externalised (Pym, 2011:1) and therefore subject to technologisation. Muñoz Martin (2016:155) goes further suggesting that such technologisation might “render a distinction between internal and external support [...] much less compelling”.

From this point, process-oriented translation research went through several phases which were informed not only by the changes in translation practice but also by more scientific methods which allowed “more systematic descriptions and accounts” of this practice (Hurtado Albir et al. 2015:6). As O’Brien duly notes, “[t]he growth in research has also come about due to the development and increased accessibility of tools and methods for measuring specific cognitive aspects of the translation task, in particular screen recording, keystroke logging and eye tracking” (2013:5). The current phase is process-oriented research associated with “empirical-experimental research [which emerged] as a result of the increasing interaction between

---

5 However, as the present thesis is concerned only with the ‘resources’ part of the ‘tool box’, when discussing translation technology only this part of the tool box will be referred to, unless otherwise stated.

6 The changes in process-oriented research related to methodological aspects are discussed in more detail in Chapter 3.
humans and machines in the translation industry” (Hurtado Albir et al. 2015:6), to which could be added the exponential growth and diversification of translation technologies and the subsequent adoption of these technologies by professional translators. Such professionalisation (Lagarde, 2009), technologisation and industrialisation of translation, together with the new methods of investigation, have had an immediate impact on the volume and direction of research in translation studies. This was duly noted by Munday, according to whom “the emergence and proliferation of new technologies have transformed translation practice and are now exerting an impact on research” (Munday, 2008:179).

This impact is reflected in the growing interest in translation technology research, both in terms of ‘surveying’ the developments and their uptake by professional translator (e.g. Fulford, 2001; Fulford & Zafra, 2004; Lagarde, 2009; Lagoudaki, 2006, 2008, Gough, 2011; Hirci, 2013; Zaretskaya et al., 2015a; Zaretskaya et al. 2015b) and in carrying out empirical-experimental research which, broadly speaking, examines the use of technology by translators and the effects it has on the translation process (Asadi & Séguinot, 2005; Podhajecka, 2006; Alves & Liparini Campos, 2009; Massey & Ehrensberger-Dow, 2010, 2011b, 2011c; Enríquez Raído, 2011; O’Brien et al., 2010; O’Brien, 2012; Valli, 2012; Zapata, 2014, 2015; Mellinger, 2014; Daems et al., 2015; Moran, 2015). As Ehrensberger-Dow & Massey sum the situation up, “[t]he increasing technologisation of the professional translation workplace has encouraged translation studies to reach beyond its traditional borders, whether to assess and develop computer-aided translation tools and machine translation or to understand the effects of technology on the translator and on the translation process” (Ehrensberger-Dow & Massey, 2014:59).

This increased research in translation technology has highlighted the complexity of the human-computer interactions that take place during the translation process in the ever more technologised translation environments. Looking at Figure 2.1, which compares by means of an example the translator’s workspace in the pre- and post-technology eras, it can easily be seen that working in an environment where several windows compete for the translator’s attention whilst they engage in the highly complex cognitive activity of translation, has rendered the workplace even more complex and thus demanding for translators, an observation also made by Ehrensberger-Dow & Massey (2014:64).
For researchers, too, the demands have increased. As the use of tools and resources made the translation process more complex, new, more comprehensive methods for capturing and analysing these complex processes were required as, externally, there is much more to be examined: the on-screen activities, the keyboard actions, the mouse movements, the logs of online activities and the trajectories and fixations of eye gaze, not to mention other data that needs to be collected to triangulate the results (see section 3.2.1. for more details). Most translators’ workspaces are now an assembly of windows, in which words, phrases or even entire sentences either pop up automatically (e.g. TM matches, Autosuggest, MT feeds, spellchecker) or in which translators can perform their own searches within their environment (e.g. concordance search, search in external resources ‘plugged in’ to their environment) (See Figure 2.1). To add more to the mix, translators often have a separate ‘Internet environment’ opened alongside their workspace7, or indeed, work with two screens simultaneously, which makes things even more complex8. A similar picture is painted by Ehrensberger-Dow & Massey (2014:70-71) in their simulation of a translator’s computer screen during a workplace study.

So, given that much of the translation process is now immersed in translation technology, it cannot be studied in isolation from it - or, indeed, from technology in general. O’Hagan points to the “deepening interrelationship between technology and the translator” (O’Hagan, 2013:507) which is further exemplified by Cronin’s reference to modern translators as “translational cyborgs who can no longer be conceived of independently of technologies with which they interact” (Cronin, 2003:112). It could therefore be argued that the increased interest in the ‘tool box’ research has brought attention to the effects translation technology might have on the translator, thus re-igniting the interest in the ‘black box’ research, or, indeed, in the intersection of the two. From the cognitive point of view, processes that were hitherto inferred or observed to be happening in translators’ heads already have or will inevitably change with the impact of translation memory software, machine translation or any future AI-assisted systems. This is something that Asadi & Séguinot (2005:526) already recognised ten years ago, referring to one of the ‘production styles’ they observed in their participants, i.e. what they call prospective thinking (see section 2.5. for more detail) as being “reminiscent of the style used in pre-computer translation”.

Therefore, it could be argued that as translation environments become ever more complex so too do the cognitive processes that need to be examined. As a matter of fact, the cognitive aspects associated with the use of technology are the subject of new research directions in cognitive translology (Muñoz Martin, 2015) and studies on the cognitive impact of the use of tools such as TM and MT have emerged (e.g. see O’Brien, 2006, 2008; Alves & Liparini Campos, 2009; Christensen, 2011; Radulescu, 2015). Furthermore, the translator’s physical environment and the changing nature of working patterns, which also largely depend on the available technology, also play an important role in placing the translation process in a wider socio-technical context. Therefore, the shift from the ‘black box’ to the ‘tool box’ emphasises the need for process research to move further beyond cognition and beyond technology alone into the realms of techno-cognition. In the context of the use of resources for translation purposes, the ‘black box’/‘tool box’ dichotomy also needs to be revisited to encompass the cognitive implications of the technologisation of

---

7 Ehrensberger-Dow & Massey (2014:69) report on a pilot study at a workplace where the combined eye-tracking and screen recording data revealed that the computer screens were too small to have the Internet browser open next to the translation memory program, forcing the translators to constantly switch between windows.

8 Speaking even more broadly about the translation process as involving all activities associated with file preparation, project management and delivery, the list of aspects of the ‘tool box’-related translation process that needs investigating expands even further and calls for examination.
the translation process, for example how translators integrate the internal resources (what they already know) with what they find in external resources.

The present study reflects this trend in its examination of the impact translation technology has on the translator and the translation process, in particular in the use of online resources seen as an integral part of this technology. As already noted, due to the scope of the present study being confined to the investigation of translators’ behaviour during their interactions with online resources, the investigation of the cognitive aspects of the technologisation of the translation process is excluded. Nevertheless, the present research is firmly grounded in the recognition that any actions translators take during the translation process is inevitably guided by internal, cognitive processes such as problem solving, decision making, filtering or prioritising.

2.2. From dictionaries to other types of resources

The technologisation of the translation process can be clearly observed when reviewing process-oriented literature from a diachronic perspective by considering the types of resources that were examined as a sole or an additional object of study. As a result of the expansion of the Internet, the democratisation of technology and, as a consequence, the wider access to a variety of resources, a pivotal change can be observed in the range of external support investigated in the process-oriented studies.

The early process studies, when investigating the use of resources, were mainly concerned with the use of printed dictionaries and encyclopaedias, as these were the only codified linguistic resources easily available to translators. For example, the ‘exclusive’ position of dictionaries can be seen in Kring’s much cited pre-Internet study (1986a) which investigated their use in eight language learners (not translators) performing a translation task. Kring’s findings that the students used a small variety of dictionaries, but for different purposes, such as comprehension, equivalents retrieval and monitoring and decision making, testify to the dominant position of dictionaries at the time of his study. Jääskeläinen (1989) also studied paper dictionary use by four translation students and found that they used a broader variety of reference aids and that some of them differed considerably in the number of items they looked up (in the present study referred to as research unit volume). Other empirical studies which partially focussed on the use of printed reference materials carried out since the 1980s include: Jääskeläinen, 1990; Kiraly, 1995; Dancette, 1997; Séguinot, 1997; Varantola, 1998 and Jensen, 1999.

Most of these studies focussed largely on the differences between novice and experienced translators, the use of mono-lingual vs. bilingual dictionaries and the frequency of dictionary use in relation to the quality of the translation, with rather predictable results. For example, Englund-Dimitrova & Jonasson’s TAP study (1999), which examined the use of dictionaries, encyclopaedias and other general reference material by professional and novice translators, highlighted some differences which were only to be expected, such as the fact that the professionals made better use of dictionaries than the novices because their knowledge of what to find in particular dictionaries was more developed due to experience.

9 Although resources such as previous translations, glossaries and consultations with experts are known to have been used centuries back as far as 4th-8th century China, 12th century Toledo School in Spain or by Bavarian monks in the 9th-11th centuries (Rogers, 2015:92-93).

10 The study revolved around dictionaries as, despite the instructions to bring any reference books that could help students with the task, most students only brought a dictionary.

11 In this study, bilingual and monolingual dictionaries as well an encyclopaedia provided by the researcher were used.
The dominant position of dictionaries could also be seen in the reportedly almost ‘intimate’ relationships developed by translators during their use, leading to an intuitive knowledge about what particular information was likely to be found in them (e.g. Englund Dimitrova & Jonasson, 1999:22; Nord, 2002:214). Such ‘relationships’ could only be ‘developed’ through long-term use of a resource that is physically bound by its covers; not one that is constantly being updated and changing as online resources now are. It could be argued that, today, this special connection has been very much weakened or disappeared, partly due to the change of medium from paper to online, but also because of the wider range of resources now available to translators.

In more recent literature, dictionaries are discussed in the context of whether they are still a dominant resource in the wake of other technologies mostly associated with the growth and diversification of Internet resources. For example, Domas White et al. (2007) claim that general dictionaries (both monolingual and bilingual) are considered to be primary sources for translators. Désilets et al. (2009) also observe that termino-lexical resources (including dictionaries) are still more widely used than resources based on text; however, the difference is not considered to be significant. This can be questioned on the basis of Massey and Ehrensberger-Dow’s survey of 110 professional translators, who declared that they use online multilingual dictionaries on a par with search engines and terminology databanks (Massey & Ehrensberger-Dow, 2011c:4, my emphasis) and Enríquez Raído’s study in which she observed a “clear trend away from the dominance of dictionaries as sources of outcomes” in her second, more challenging embedding task (Enríquez Raído, 2011:458). A noteworthy observation with regard to dictionary use made by Désilets et al. (2009:np) is that the “subjects overwhelmingly used more bilingual resources than unilingual ones and the difference was highly significant”. This supports the findings of early process studies and recent studies such as Torres Domínguez’s (2012) survey of 509 translation students and professionals, Durán-Muñoz’s (2010) survey of 402 translation professionals and the Global Survey of 540 translation professionals discussed in the present work.

Moving on from the focus on dictionaries in process-oriented research towards a more comprehensive overview of the external resources currently found in a translator’s ‘tool box’, an extensive list of resources under investigation can be compiled from recent empirical studies (Pavlović, 2007; Désilets et al., 2009; Domas White et al., 2007; Massey & Ehrensberger-Dow, 2011c; Enríquez Raído, 2011). The resources reported as being used by translators include printed and online dictionaries, encyclopaedias, parallel texts, search engines, terminology databases, lexicons, translation memories, bitexts, online documents, rules for stemming, books on usage, annotated dictionaries, journals, almanacs, cultural encyclopaedias, search engines, online encyclopaedias, discussion fora, academic journals and a host of academic, organisational and commercial websites. Furthermore, Zapata (2015:149) reports that even when post-editing, translators use term banks, dictionaries, corpora, concordancers and search engines. The present work seeks to provide an even more detailed account of resource use by considering self-declared and observed data from a language-independent sample of experienced, professional translators, an aspect of studies on

---

12 Although translators might be developing different, more temporary or explorative types of ‘relationships’ with their online resources.

13 Although the present study is not directly concerned with post-editing but with how translators use online resources during their translation processes, references to post-editing are made throughout this thesis as post-editing practices were observed to be part of these processes in the present study.
resource use that has not been sufficiently addressed yet, especially from the information-seeking perspective.

A broader perspective in relation to the examination of the translator’s ‘tool box’ is adopted by researchers at the Copenhagen Business School where recent studies have focussed on optimising “user interface design and assistive technology” (Angelone et al., 2016:50), especially machine translation. Two studies that were recently conducted and are concerned with the use of external resources during translation and post-editing are of particular relevance to the present study. Firstly, Daems, et al. (2015) report on their study of ten MA students of translation in which the use of resources was examined during the translation and post-editing of eight texts using a translator’s workbench and the logging tool CASMACAT14. Daems et al. studied eye tracking, key stroke and web log data related to the online activities and then calculated the number of times a resource was consulted and the time spent in that resource during the session. They found that “significantly more time is spent in external resources in human translation, compared to post-editing”, which increases the overall translation time (Daems et al., 2015:122), whilst the types of resources used in both processes, as well as overall quality (measured in terms of acceptability and adequacy) were comparable (Daems et al., 2015:121-130). The authors conclude that “even though the resources might help translators solve translation problems, this goes at the cost of overall productivity” (Daems et al., 2015:134). In future Daems et al. propose to examine professional translators and, with relevance to the present study, to look at ‘between participants’ differences (Daems et al., 2015:134-135).

Secondly, taking a slight diversion into the realm of exclusive post-editing, Zapata (2015) reports on a pilot study in which he examined the use of external resources during two types of post editing processes. Seven professional translators post-edited two specialised medical texts using resources such as BiConc (a prototype bilingual concordancer) built into CASMACAT as well as other online resources. They performed the task under two conditions: traditional post-editing and post-editing through interactive translation prediction with online learning. The data was logged in CASMACAT and screen recorded and additional eye tracking data was also collected. As mentioned above, it appears that even during the post-editing process the use of external resources is significant (Zapata, 2015:149) and translators’ interactions with the information encountered in these resources are complex. Further systematic examination of the use of resources, especially in the context of using assistive technologies such as translation memory or machine translation is needed.

To take full account of the impact of the multitude of resources available to translators today and of the often multimodal nature of these resources, a strong new direction to consider in future ‘tool box’ research is multimodality in the translation process (Mees, 2013; Zapata, 2014; Mesa-Lao, 2014). In the context of human-computer interaction, efforts directed at “finding alternatives to traditional keyboard-and-mouse input [...] such as voice- and touch-enabled multimodal interfaces” are currently being undertaken in order to introduce these alternatives into the translation process (Zapata, 2014:339). In view of the ever-increasing complexity of the translation environments discussed earlier, research is needed not only to monitor the adoption of these new modalities by translators, but also the impact of these new technologies on the translation process, and, of particular interest to the scope of the present study, on the use of external resources.

---

14 http://www.casmacat.eu/
Having discussed the changes that can be observed in the scope of the investigated resources, it can be argued that the pre-Internet studies investigating the use of resources need to be interpreted in the context of a translation process that was very different indeed from what we observe today; one where dictionaries occupied a central position in the absence of the many resources that are now available. And yet, this is not always acknowledged in process-oriented literature and the research findings of the pre-Internet era seem to be often treated on a par with the more recent ones, as if no change had taken place with regard to the availability or indeed, accessibility of the resources during the studied translation process.

Furthermore, the findings related to the patterns of resource use can only be interpreted in the context of how these resources were made available during these studies, i.e. they were often supplied by the researchers themselves\(^{15}\). Today, the ‘resource playing field’ is completely different as those participating in research studies have equal and unrestricted access to a host of quality, online resources. It is also possible that in time-restricted studies, the participants consulted resources less frequently (compared to today) purely because of the time it would have taken them to search through a paper copy.

In fact, Luukkainen's study, designed to compare the translation process used with and without resources\(^ {16}\), did indeed conclude that the use of reference material negatively impacts on the time spent on translation by slowing it down and limiting translators’ creativity (Luukkainen, 1996:71). Similarly, House carried out a comparable experiment in which ten advanced language students translated with and without access to reference material. She found that some students (the ‘low-risk-takers’) relied on reference material more than others (the ‘high-risk-takers’) and links the performance of the ‘high-risk-takers’ with, for example, higher levels of confidence, fluency and creativity (House, 2000:155-158). A study by Livbjerg and Mees (1999) was also carried out to compare the translation process with and without the use of reference material. In their research design, however, the use of reference material was offered in a second task to correct the text translated without any recourse to reference works in the first task. Livbjerg & Mees reported only marginal gains from using dictionaries and the study was repeated with some changes three years later leading to similar conclusions that the overall value of dictionaries is rather limited, despite the fact that they were seen to contribute to the quality of the translations.

As with other early studies examining the use of resources by translators, the conclusions from Luukkainen, House and Livbjerg & Mees's studies need to be interpreted in the context of their pre-Internet time, when the reliance on a limited pool of resources requiring a different, possibly more time-consuming and less ergonomic type of interaction than that now offered by online resources, could have been perceived as hindering the translation process, thus interrupting, “the flow of thought”, as House puts it (2000:156). It could also be argued that translators were somehow restricted to the solutions available in the paper resources in which they invested heavily, having no easy access to alternative sources as is the case today\(^ {17}\). Therefore, the move from paper to electronic resources in response to the rapid expansion of the

\(^{15}\) This argument could also be made against carrying out studies in laboratory settings. As Nord’s study (2009) shows, more types of resources were observed to be used when translators were examined in their natural environments.

\(^{16}\) A range of paper resources such as bilingual dictionaries, monolingual dictionaries, encyclopaedias and parallel texts were examined.

\(^{17}\) For example, owing to cost implications, they might have had to make choices as to which particular dictionary to purchase thus missing out on information contained in alternative dictionaries.
Internet suggests a corresponding shift in the way the resources are consulted by translators and integrated in the translation process. Although on the face of it the interactions between the source text and online resources appear rather fragmented, the changes in translation practice (the move from paper to online and general sharp increase in available resources) raise the question of what exactly these changes look like.

It could be posited that online resources are more responsive to the needs of translators and, therefore, their role in the translation process might be perceived more favourably than that of their paper counterparts despite any disruptive effects they might have on this process. Now, thanks to the speed of access and the diversity of available resources on the Internet, online resources supporting the translation process have become the norm in everyday translation practice. This, in turn, has markedly shifted the research perspective by considering the presence of technology in the translation process as a given rather than an option. Therefore, the more recent process-oriented research, which has seen translators increasingly using the Internet and resources becoming available online, has broadened the range of resources being investigated within the translator’s ‘tool box’, extending it far beyond the scope of dictionaries. Thus, in process-oriented research, the focus has shifted away from the questioning of the relevance and usefulness of resources in the translation process towards investigations into the adoption of new technologies and observations of their actual use, as well as how they can be improved.

The present research concentrates on such investigations into examining the part of the tool box which encompasses online resources and looking at their use by professional translators alongside any other technologies they have as part of their usual translation environment, such as translation memory or machine translation and in any mode (translation or post-editing) chosen by the participants. It is hoped that the present account of the range and availability of online resources and the patterns of their use by professional translators will provide a much needed insight into their actual practices, thus answering O’Hagan’s call for a better “understanding of the nature of mediation by technology to uncover the complex relationship being formed between the translator and the tool” (O’Hagan, 2013:514), extending it to the relationship between the translator and the resource. As will be argued in Chapter 6, in the light of the rapidly growing number and diversity of available resources, translators might become more selective about the type of resources they use. Such selectiveness can manifest itself in a certain type of resource behaviour, i.e. the type of resources that are dominant in each translator’s online interactions and the way in which these resources are accessed. It is therefore argued that as the tool box undergoes even further changes in the future, more research will be needed to see how translators interact with new technologies, new resources, new processes and new types of content and the way it is delivered.

2.3. A growing interest in translators’ online information behaviour

As discussed in the previous section, the technologisation of the translation process and the emergence of the World Wide Web as a “virtually ubiquitous information environment” (Bawden & Robinson, 2011:130) have had a profound impact on the kinds of resources that are used for translation-oriented research.

---

18 Cisco report on the steep Internet growth curve: “In 1984, total global Internet traffic was 15 Gigabytes per month. By 2014, the average Internet traffic per user was 15 Gigabytes per month. Each one of us generates more traffic than crossed the entire global Internet 30 years ago” (Sumits, 2015).
However, these changes necessitate the examination of how these resources are accessed in this online environment and this requires a perspective from information behaviour studies.

*Information behaviour* (IB), as defined by Wilson (2000:49), refers to “the totality of human behaviour in relation to sources and channels of information, including both active and passive information-seeking, and information use”. It includes those activities a person may engage in after identifying their own needs for information, including searching for such information, using it or transferring it (Wilson, 1999:249). In the present work *information behaviour* is perceived from an anthropological point of view and understood to be “an instinctive human mechanism, developed and expressed in interaction with the environment” (Spink & Heinström, 2011:5). This stance allows the researcher to consider the information behaviour of translators not only in relation to variables such as age, experience, domain knowledge or the attributes of the task being undertaken, but also as an expression of their natural, individual predispositions. Following Enríquez Raído’s (2011) adoption of Wilson’s nested view of information behaviour, the present study also draws on this model (see Figure 2.2). In this model, *information seeking* includes a variety of methods employed to discover and gain access to information resources in the context of translation - paper, electronic, online resources and consulting experts, while *information searching* is “particularly concerned with the interactions between information user […] and computer-based information systems” (Wilson, 1999).

![Figure 2.2 A NESTED VIEW OF INFORMATION BEHAVIOUR (Wilson, 1999)](image)

However, since external resources used by translators are now predominantly ‘computer based’, in the present study information seeking and information searching are conflated and a term *online information behaviour* used as an umbrella term. However, in some contexts requiring a wider perspective that includes non-computer based resources, it is still necessary to use the term *information seeking*. This particularly applies when discussing research activities as distinguished from translation activities.

Looking at the literature in the area of information behaviour it can be seen that the main focus is an occupational one (Case, 2012:286). The groups studied include scientists and engineers, social scientists, humanities scholars, health care providers, managers, journalists, lawyers and even farmers, artists or clergy19 (Case, 2012:285), but no studies relating to translators are mentioned in a recent volume of some

---

19 See Case (2012:288-289) for a comprehensive list of literature related to studying information seeking, needs or behaviour of these groups.
490 pages devoted to research in information seeking, needs and behaviour (Case, 2012). Furthermore, although the information behaviour of various groups of people defined by profession, occupation, educational level, role (e.g. voter, patient etc.), demographic group or even hobby has been studied rather extensively (Bawden & Robinson, 2011:1), studies examining the individual differences within those groups are less frequent, even though there is a lot of within-group variation in their information behaviour. As Line observes, “doctors, graduate students in humanities, or job-seekers [...] do not form homogeneous groups” and their individual differences do not have to be related to their age, experience or educational level (ibid). In the quote below, Line (1998:223) summarises the fact that all individuals seek information in different ways and that the way we look for information is part of our personalities.

There is no such thing as a homogeneous body of information users. A group of people with the same level of education may be working on the same subject at the same time, but they will use information in different ways. Some information will obviously be relevant to all of them, some to only one or two. Some of the group will scan and absorb vast quantities of material, others will read much less. Some will be content with abstracts of many articles, others will not. Some will prefer oral channels, some will prefer to see information in print [...], some always prefer to ask other people, some prefer to avoid people whenever they can. Some like browsing, some find it wasteful in time and effort.

(Line, 1998:223)

Therefore, as Domas White et al.(2008:577) rightly put it, “the time is opportune to study information needs and behaviors of professional translators” and the present study is a contribution to this area of study. Moreover, to paraphrase Domas White et al., it could also be said that the time is opportune to study not only common tendencies in these behaviours, but also the individual differences between them. Although people look for information in different ways, it is often possible to recognise certain patterns within the group, leading to the identification of sub-groups and capturing certain similarities and tendencies within those sub-groups, especially in larger samples.

Despite the growing body of research in the area of information behaviour, little is yet known about how particular groups of users interact with web-based resources. Lewandowski laments that research in the area of web searching is still in its infancy, and “[w]hile technical development is fast and lots of research is published in that area, with regard to gaining a deeper understanding of the user, the searching process, and the societal impact of search engines (not to mention the combination of these) there is still only limited understanding” (Lewandowski, 2012:2). Based on the available information behaviour studies the user classifications these studies propose, the following review will discuss those of the most relevance to the present study of professional translators and their online information behaviour.

One of the studies of relevance to the present work is that of Broder (2002) who carried out a survey and query log study in order to come up with a taxonomy of web searches. To put things in context, however, this study was carried out thirteen years ago based on the now obsolete AltaVista search engine. Nevertheless, Broder recognised that in the context of web searching, not all needs behind a search query are indeed informational in nature; they can also be based on the need to reach a particular website or to find a source from which a transaction such as a purchase or a download can be performed. Thus he distinguished between three types of queries: navigational, where the immediate intent is to reach a particular site, informational, where the intent is to acquire information, and transactional, where the intent is to perform some web-mediated activity (Broder, 2002:5). Based on the log query analysis (not the survey that was also carried out as part of the study), 20% of the queries were navigational, 48% were
informational and 30% were transactional. Later studies show a growing dominance of informational queries. For example, Rose & Levinson (2004) confirm the dominant presence of informational queries in their study. Six years later, Jansen, Booth, & Spink (2008) used Broder’s three-level model of classification of web searching and, based on data gathered from several hundred thousand users (over 1.5 million queries) automatically classified by purpose built software, they found, with 75% accuracy, that most of the queries (80%) were informational in nature. It is highly likely that these statistics have shifted again since 2008 as the Internet has become even more focussed on the transactional element associated with multimedia content, online shopping and social interactions. Nevertheless, Broder’s classification is still being referred to today in information behaviour literature and has a particular importance for the present study (see section 6.1.3.). However, in translator online behaviour both navigational and informational queries can be perceived as looking for information; the difference is in the way the information is obtained - through a search engine query ( informational) or through accessing a website in a direct fashion (navigational).

Based on two predominant cognition and problem-solving styles derived from models of cognitive processing, the convergent-rational and divergent-intuitive (Edwards, 2003), Heinström investigated the information seeking patterns of 305 MA students in relation to their discipline of study, their approach to studying and personality. The analytical framework is based on a dimension most commonly found in literature, i.e. that of “broadness versus specificity” (Heinström, 2006a:1440). Figure 2.3 below represents the main findings, namely three information seeking styles – broad scanning, fast surfing and deep diving.

![Figure 2.3 DIMENSION OF EXPLORATION AND SPECIFICITY RELATED TO INFORMATION SEEKING STYLE, STUDY APPROACH AND PERSONALITY - Figure adapted from Heinström (2006a:1449)](image)

Heinström (2006a:1449) suggests that further research is needed in a variety of contexts, especially “related to work tasks or everyday life information seeking” (Heinström, 2006a:1449). As will be seen later, the present research addresses this call and some of the characteristics of the information seeking styles found by Heinström can be recognised in the online information behaviour of professional translators as manifested in their translation-oriented research activities.
Other characteristic behaviours identified in a large-scale study of the Google generation\textsuperscript{20} (Rowlands et al., 2008) are associated with shallow, skimming activities, short sessions, power browsing, squirrelling, i.e. downloading information and cross-checking behaviours from various websites. However Rowlands et al. argue that these search behaviours are not only associated with younger users, but “from undergraduates to professors, people exhibit a strong tendency towards a shallow, horizontal, ‘flicking’ behaviour in digital libraries. Power browsing and viewing appear the norm for all” (2008:300). As will be seen later, this finding is of relevance to the present study.

Furthermore, ‘information encountering’ and the attendant notion of ‘incidental information acquisition’ (Elderez, 1997; Bawden & Robinson, 2011) are amongst other concepts that are relevant to the present study. These concepts relate to a behaviour that is in contrast to the purposive search strategies discussed earlier and in the present research has been linked to one of the proposed macro-research strategies, in which information encountering is very much a part of the interactions with parallel texts during the translation-oriented research process, where potentially useful information is identified indirectly for future use in the translation.

However, most importantly for the present work, information behaviour is understood to be not just a phenomenon associated with technology and the so called ‘information society’ we now live in, but an inherently human mechanism that evolved into a “genetically favoured trait” (Spink & Heinström, 2011:3). As such, it also takes into consideration the context, culture and environment in which it is shaped (Spink & Heinström, 2011:4). From this perspective, it is possible to conceive that professional translators, just as any other group, display certain characteristic patterns of information behaviour which are tailored to their specific needs as translators, but also, within the confines of translator-information behaviour, they exhibit their own individual, innate patterns, i.e. styles. Spink & Heinström argue that assimilating such a perspective of information behaviour will help further the understanding of information behaviour and help us “move beyond the contemporary focus on information systems and processes within the field” (Spink & Heinström, 2011:4). Furthermore, “linking information behaviour to other instinctive genetic dispositions, such as personality differences and language, we can further increase our understanding of information behaviour patterns and styles” (ibid.).

The present study is an attempt to further such knowledge and it taps into the important and growing trend in ‘tool box’-oriented translation research relating to the examination of the use of external resources from the broader, interdisciplinary perspective of information behaviour, and consequently, the focus on translators’ online research activities. Whilst it was only natural to examine the use of paper resources in the pre-Internet era (see section 2.1.), in the era of digital humanism (Cronin, 2013:7) in which “[t]he working practices of translators have been changed beyond recognition in terms of the access to many different kinds of knowledge that are afforded by the infrastructure of the internet”, studying how translators find information online seems ever more timely. And yet, despite the fact that translators are seen as “information users, processors and producers” (Pinto & Sales, 2008:413) who nowadays use, process and produce this information mostly immersed in online environments, until recently “translation studies research [...] has rather neglected this important aspect of translator behaviour” (Enríquez Raido,

\textsuperscript{20}Google generation refers to young people “born after 1993, growing up in the world dominated by the Internet” (Rowlands et al., 2008:292).
2014:Preface) and the “technical and information behaviour of translators has only recently emerged as a significant area of research in Translation Studies” (Massey & Ehrensberger-Dow, 2011c:9). Indeed, since Domas White et al.’s call to study translators’ information behaviour in 2008, some researchers have taken an interest in this area and several studies have emerged. Due to the multi-dimensionality of translator online activities and the different professional settings in which they occur, these studies vary in many respects and examine various aspects of these activities.

For example, Valli looks at user behaviour in relation to concordancers, i.e. tools “commonly used by translators to find a target language version of the source text portion they entered as a search string” (Valli, 2012:1). She examines search strategies, types of interactions and recurrent search patterns in concordancers as well as looking into how translators use concordancers by analysing the settings and filters they apply to searches. The study is based on a large collection of search logs from the EU concordancing tool Euramis. Valli sees each concordance search as “a complex event where translation skills, search strategies, computer skills and translation-oriented Information Retrieval (IR) are combined” (Valli, 2012:2). Insights from such unique studies are extremely valuable not only from the point of view of the development of translation tools, but generally as evidence that translation activity has become essentially an online phenomenon, and, although search behaviour during translation activity might differ from general online searches in some respects, cognitively, it shares basic activities such as “developing a search strategy, balancing recall and precision or engaging with results assessment” (Valli, 2012: 2).

Enríquez Raído’s (2011) pedagogically motivated research into the web searching behaviour of six student translators and two professional translators also approaches problem-solving in the translation process from the information behaviour perspective and greatly contributes to the understanding of how translators use online resources during the translation process as part of their documentary research. Based on data from her multi-method study, she identified two web searching behaviour styles, linking each to the level of expertise and declarative knowledge of translation of the individual participants. The ‘shallow’ style associated with horizontal “checking and comparing type of search behaviour [...] [involving] easy, fast, and more or less cursory visits to a few selected websites” was found to be characteristic of student behaviour (Enríquez Raído, 2014:139-140). A deeper and wider, ‘interactionistic’ style involving a more browsing type of search, i.e. following internal links, performing site queries and spending more time reading to acquire background information was attributable to the more experienced translators (ibid.). Individual aspects of the subject’s search behaviour are described in much detail based on a ‘case study’ approach; but the overall analysis is focussed on the differences resulting from the increased task complexity rather than from the differences between the participants per se. Therefore the present research is, to my knowledge, the first of its kind to address the individual differences within a group of translators and to attempt to group these differences into ‘types’. As will be seen in Chapter 7, the two

22 Two translators of varying levels of expertise took part in the pilot study.
23 The task complexity was a factor that did influence the search behaviour of the participants in Enríquez Raído’s study, for example by lengthening the search sessions or through the use of more targeted and complex query statements; but this is to be expected, as the Pilot Study of the present study also shows. However, the interesting insight resulting from the increase in task complexity is the change in the type of resources used. Enríquez Raído reports that one participant’s behaviour changed leading to more frequent consultation of text-based resources in the second, more demanding task (2011:473). This change is attributed to the learning curve, but could have implications for the present study, as such change could not be observed due to the fact that only one translation task was performed.
search styles identified by Enríquez Raído can be incorporated into a more comprehensive narrative regarding translator research styles.

Enríquez Raído’s study further stresses the importance of exposing students as early as possible to the diversity of online resources available to them, teaching them to recognise their information needs based on the translation problems they encounter and to select appropriate resources based on these needs. Also, when it comes to the use of search engines, two areas of training have become apparent: “a) knowledge about search engine performance for data retrieval as opposed to information retrieval, and b) the use of search engine features to efficiently construct (advanced) queries for finding relevant information on the web” (Enríquez Raído, 2014:182).

Inspired by Enríquez Raído’s 2011 study, Volanen examined the online information behaviour of five professional translators who carried out their translation task under laboratory conditions, with the general research goal of shedding light “on the information seeking behaviour of professional translators in terms of online resources” (Volanen, 2015:43). The study was carried out using resources exclusively available on the Internet and examined the types of online resources translators used and their information-seeking actions. Additionally, translators retrospectively commented on their perceived level of performance as well as on how they had fared using exclusively online resources. Comparing her results to those of Enríquez Raído, Volanen concludes that professional translators spend less time researching individual translation problems (called, in her study, IS contexts) and that they are “more efficient in their information seeking in general, because they possess more knowledge on the ST topic or merely on the available online resources” (Volanen, 2015:63).

Massey and Ehrensberger-Dow were also amongst the first researchers to start a systematic examination of the translation-centred information behaviour of professional translators through their project ‘Translation Tools in the Workplace’. In their multi-method study they examined the “self-reported resources use and information behaviour of professional translators” as well as the observed “tool use and research practices in the workplace” (Massey & Ehrensberger-Dow, 2011c:3). These investigations were carried out with a strong ergonomic focus showing “how ergonomic issues can adversely affect both the speed and efficiency of the translation process and, ultimately, the success of the translation product” (Massey & Ehrensberger-Dow, 2011c:4). These ergonomic issues were identified as problems with multiple window management, various difficulties with the interface of online resources such as online dictionaries, search engines or online encyclopaedias (Massey & Ehrensberger-Dow, 2011c:6) and a phenomenon the authors call “looking but not seeing” (Massey & Ehrensberger-Dow, 2011c:8) which could be partly related to online reading habits or problems with the way search results are presented (ibid.). Broadly speaking, the ergonomic problems identified included “inefficient resource and desktop management [and] deficient knowledge of (automated) features and ineffective interaction with user interfaces” (Massey & Ehrensberger-Dow, 2011c:9).

A complementary endeavour is the research conducted by Pinto et al. in their InfoLiTrans project (Pinto & Sales, 2008; Pinto et al., 2010; Sales & Pinto, 2011; Pinto et al., 2014;) which concentrates on developing information literacy instruction for students, teachers and academics and professionals. In the first phase Pinto & Sales (2008) concentrated on translation and interpreting students and aimed at obtaining “a diagnosis of information competences” by means of implementing a purpose-built test across seventeen universities in Spain. Out of the four categories focussing on different information management competences (information search, information assessment, information treatment and communication and
dissemination of information), the weakest point proved to be information searching (Pinto & Sales, 2008:90). Other areas of improvement were identified as reading (perhaps associated with the ‘looking but not seeing’ phenomenon mentioned earlier), theoretical and practical knowledge of available online resources and primary information sources. With regard to the improvement of search techniques, Pinto & Sales recommend using real examples from real translation experiences to reinforce good translator research practices (2008:90). The second phase of Sales & Pinto’s research (2011), in which they report on data obtained in a survey conducted in 2007-2008, extended to identifying the information needs of 101 translation and interpretation professionals as they fully recognised the fact that translators have to be able to locate, retrieve, handle and disseminate the information, and to manipulate “the extraordinary new resources that information and telecommunications technologies have made available for [their] work” (Sales & Pinto, 2011:246). Durán-Muñoz’s (2010) study also examines translators’ needs, but this research did not stem from the recognition that translation is essentially an online activity, but rather from the pragmatic point of view that the “majority of resources currently available are of little use to translators” (Durán-Muñoz, 2010:1).

Mutta et al. (2014) studied information seeking in the slightly different context of language learners (L1 and L2) and assessed students’ digital literacy based on “the critical assessment and use of digital technologies and the competencies in digital communication and discourse” (Mutta et al.,2014:1). The task was to search for information related to ‘factual tasks’ of varying degrees of difficulty, ranging from simple questions such as ‘What is the height of the tower of Pisa?’ to impossible questions such as ‘What is the frequency of the radio station ‘Santé?’24. Because one subgroup had to search in their L225, translation was considered to be a “central part of the information-seeking procedure” (Mutta et al.,2014:4). They found - unsurprisingly - that students spent less time searching for information in their native language and that more queries were formulated whilst searching for information in their non-native language. With particular importance for the present study, the results also showed that students developed individual seeking paths and strategies to complete the tasks and that “the main reason for differences is individual search behaviour” (Mutta et al., 2014:14).

Despite the growing interest in studying information behaviour in relation to the translation process, researchers emphasise the fact that much more needs to be done in this area. Valli, for example, laments that “there is hardly any data available that takes into account the new types of online linguistic resources and investigates user needs and translators’ online search behaviour” (Valli, 2012:2). Zapata (2015:151) also calls for greater attention to be paid “to the study of translators’ interaction with (digital/Internet-based) information” and Daems et al. see the need to develop a “better understanding of the usage of external resources during translation and post-editing […] to obtain a more profound insight into successful problem-solving strategies” (Daems et al., 2015:115). The present study attempts to provide a modest contribution towards at least partially closing this gap. However, the rapid transformation of the translator’s ‘tool box’, the developments of Internet technology and the advancements in artificial intelligence call for an outlook beyond tools and resources and beyond the study of translator-computer interaction (TCI) (O’Brien, 2012) to translator-information interaction (TII) (Zapata, 2015).

24 Presumably this radio station does not exist.
25 This group of Finnish students also studied translation into French.
2.4. From laboratory to workplace and novice to expert

As mentioned in section 2.1., early process research was confined mostly to the classroom or the laboratory and the translation process was studied using mostly language learners, translation students and occasionally more experienced translators who happened to work at an institution where studies were being carried out. However, when investigating translation processes it is important to bear in mind that translation is also a profession and that translators work in a variety of settings: as freelancers at home, as in-house translators in translation agencies, in companies and government institutions. Therefore, as Vienne observes, we need to “focus on real-world professional translation as enabling intercultural communication instead of translation as a means of linguistic comparison or language learning” (Vienne, 1994, original emphasis, in: Risku, 2002:524). Although examining the translation process in a classroom setting can bring many important pedagogical insights, Ehrensberger-Dow & Massey argue that “investigating translation processes becomes truly relevant to translation competence and practice when the processes reflect actual practices of working translators, not artefacts of experimental settings and tasks” (Ehrensberger-Dow & Massey, 2015:11). Hence, interest in studying expert behaviour in natural settings has been growing within process-oriented translation research. Furthermore, a new perspective of the translation environment as a complex system involving interactions with people, physical environment and cultural artefacts (Risku, 2002) is required in order to fully grasp the changing nature of the translation process as a “highly complex problem-solving process embedded in social and physical environments” (Risku, 2002:525). As Massey & Ehrensberger-Dow (2010:128) observe, “[w]hile developments in software applications and business processes in translation service companies reflect the changes taking place in the workplace, little research has been done into how they affect the cognitive processes and working practices of language professionals”. The present research is also very much motivated by the need to study professionals using technology in their natural environment, with particular focus on their use of online resources in their working practice. Moreover, it is the first of its kind to study freelance professional translators carrying out a task in their own workplaces as the trend towards moving research studies from the laboratory to the workplace naturally entails a shift in the profile of the participants, from students to professionals.

Most of the early process studies were carried out on language learners or students of translation, as we have seen, although some researchers, e.g. Luukkainen (1996), considered fourth-year students as experienced translators and for Tirkkonen-Condit (1989) fifth year students were considered ‘professional’. Varantola (1998) also used students under laboratory conditions despite the fact that her study concentrates on the needs of professional translators. Much research has also been carried out focussing on the differences between students/novice translators and professional translators, but professional translators have only recently started to become the focus of attention in process-oriented translation research. Therefore, the object of study in all the literature discussed below is the professional translator and the processes studied are closer to expert behaviour.

One of the first studies to examine professional translators in their workplace was Nord’s study of thirteen professional translators carried out in 1997, published in German as a PhD thesis in 2002 and retrospectively reported on in 2009 (Nord, 2009). In this pioneering study Nord observed the translation process of professionals in “authentic translation situations, [...] translating texts of their own choice arising

---

26 Although non-professional translation has always existed and new forms of online community translation or crowd sourced translation, where translation is not carried out in professional settings, have also emerged with web 2.0.
in their daily routine, using aids normally available to them” (Nord, 2009:203). Nord obtained detailed statistics on six parameters, including the frequency of consultations, the types of resources used, the query types indicating the type of information sought (e.g. meaning, equivalent, use, grammar, collocation, alternative etc.) and the success of the actions. Amongst other findings, Nord emphasised the importance of resources in translators’ work. She lists several different types of resources used by translators such as different kinds of dictionaries (monolingual, bilingual, synonyms), encyclopaedias, personal databases, atlases/chronicles, auxiliary texts and ‘persons’ (i.e. human consultations). She also points out that, on average, the translation process was interrupted every 3.5 minutes so that reference material could be consulted (Nord, 2009: 214-215). Nord’s study is of great importance to the present work as it provides the only basis of quantitative comparisons of resource use by freelance professional translators in the pre- and post-Internet eras.

Another early study of professionals in the workplace was carried out by Lauffer (2002). In this multi-method study, Lauffer observed a junior and a senior translator at Toyota Canada, and another translator working at the Ontario Government’s Translation Service. The translators were “in their natural work environment with the tools they were accustomed to using” and had their “computers, databases, and electronic and traditional reference materials” (Lauffer, 2002:60). Lauffer observed three strategies within the translation process: understanding and reasoning, searching and revising. However, in line with Séguinot’s (2000:146) observations of some ‘iterativeness’, non-linearity and parallel processing in the translation process, Lauffer admits that “[t]hese strategies are not clear cut or straightforward, and they overlap and reoccur throughout the translation process” (2002:68). As will be seen later, these findings are highly relevant to the present study.

In another workplace study, Hébert-Malloch (2004) observed one translator’s work at the Federal Government’s Translation Bureau in Canada for 26 days, generating 152 hours of videotape material that showed what was happening on his computer screen27. Although the article was published in 2004, the study was actually carried out much earlier, around 1998/199928 and so it can be assumed that not many online resources were in existence at the time as there is no mention of online consultations in the analysis. The study provides some calculations as to the number of words translated, speed of translation and time spent on translation. More importantly for the present study, the researcher provides some statistics on pauses which are divided into internal (happening during the interactions with the source text and lasting between 3” to 39’) and external (happening between texts or during the lunch hour and lasting between 3’ and 6h 29’). The internal pauses are of relevance for the present study as they could relate to the use of resources. Unfortunately, the figures do not seem to add up29 and therefore, it is difficult to draw any conclusions from this study other than that only 41% of all observed time was actually spent on interactions with the texts and that not all pauses were attributable to research activities. It is interesting

27 The method was not screen recording, but the researcher “‘hooked up’ the translator’s workstation to a videorecorder and recorded his work directly onto a videotape after having installed a special video card in his computer” (Hébert-Malloch, 2004)
28 It is not mentioned in the paper when the study was carried out; however, the author says that the article is the summary of an MA thesis defended in 1999.
29 A total of 151h 47‘ (9088’) was recorded in 26 days, which averages 5.8 hrs a day. Of the total recording time, 89h 14’ was spent on pauses (5348’), which accounted for 59% of the total recorded time. As mentioned, pauses were divided into internal and external and the researcher attempted to calculate out the external pauses in order to get a better picture of the translation process itself without lunch breaks etc. However, the figure of 5409’ (90h 15’) given to indicate the time that passed “[d]iscounting external pauses” (Hébert-Malloch, 2004:975) is actually greater than the time that included all pauses (89h 14’).
to see, however, how difficult any quantitative measuring of the process was before the availability of tools such as Translog\textsuperscript{30}, or recently tools designed to measure productivity in relation to MT post-editing such as iOmegaT (Moran et al., 2014).

In another, more recent workplace study, Karamanis et al. (2011) investigated localisation activities in the localisation division of a software company and two Language Service Providers (LSPs), a large one employing 150 in-house translators and a small one employing four in-house translators. Karamanis et al. provide interesting, hands-on insights related to the nature of collaborative efforts in the translation process as well as the relationships between in-house and freelance translators during these processes and, with particular relevance to the present study, they highlight the nature and the extent of research activities that are carried out unbeknown to the client. The examination of the realities of the workplace can reveal much needed detail about how the translation process is embedded in other processes and procedures. For example, Karamanis et al. provide an interesting perspective on the relationship between in-house translators and external freelancers working on the same project. They report that “[f]reelancers are trusted less because they are not full participants to the way the work is done within the LSP”. They are reportedly also given less challenging parts to translate and are reviewed more thoroughly than the in-house staff (2011:9). Further studies related to translators working in translation agencies are reviewed in Kuznik & Verd (2010) and LeBlanc (2013), in European institutions, see Koskinen (2008) and Valli (2011, 2012), and in companies, see Asadi & Séguinot (2005).

The many demands of workplace research were highlighted by Ehrensberger-Dow (2014) based on her recent, naturalistic workplace study carried out in a large Swiss LSP as part of the project ‘Capturing Translation Processes’\textsuperscript{31}. In a detailed account of the challenges encountered during her research, Ehrensberger-Dow emphasises the difficulties associated with conducting workplace research (as opposed to a lab-based experiment) and the many compromises that had to be negotiated during the process. This practical guide through the challenges encountered whilst conducting research in a translation company (LSP) provides an invaluable insight for future researchers who wish to embark upon workplace research, making the point that the most important aspect to consider in this type of research is flexibility (Ehrensberger-Dow, 2014:361). Whilst the difficulties identified might certainly apply to many workplace studies such as those carried out in institutions and translators working as in-house translators in companies and LSPs, the demands posed by workplace studies of freelance translators are of a somewhat different nature.

In a workplace study such as the present one, where freelance professional translators carry out their task in their own place of work, the challenges are slightly different. They might include, for example, finding the required number of suitable candidates willing to take responsibility for data collection, designing the study to carry out Internet-mediated research (e.g. suitable software for data collection, installing necessary software for data collection on their computers), data security and ethical considerations. They can also relate to assessing the level of ‘professionalism’\textsuperscript{32} of the participants, controlling the delivery of the task, or even getting the participants to cooperate in the first place. Calling for more ecological investigations into the natural translation processes of professional translators, “including professional use of reference works”, Schneider (2001:158) notes that “[a]dmittedly, it is no trivial matter to persuade

\textsuperscript{30} https://sites.google.com/site/centretranslationinnovation/translog-ii/

\textsuperscript{31} Earlier studies as part of the same project concentrate on the ‘workplace’ of students and are therefore not part of this analysis.

\textsuperscript{32} This is often mediated by using the length of their experience in full time self-employment as a translator.
professional translators to get involved in scholarly experiments [...] it is much easier to work with student populations instead, under laboratory conditions, particularly with your own students” (Schneider, 2001).

Indeed, the challenges associated with the recruitment of suitable candidates cannot be underestimated. Even by applying filters such as length of experience as a full-time translator, relevant education and interest in continuing professional development, it is impossible to avoid other important variables that might have an influence on the results of the study. One example of such inconspicuous variables could be the strata that exist within the community of professional translators and the services they offer. Commenting on a blog post (2015), a well-known professional translator Chris Durban points out that there is a “spectrum ranging from what we’ve termed ‘bulk’ (large volume, often high churn, low bar for quality, low perceived risk, low prices, poor working conditions) [...] to market segments where clients are savvier, more demanding (for both subject-matter knowledge and writing style), more available for questions & assistance [...], and where budgets are also far higher, even though turnarounds may still be tight”. Such differences would most likely be an important variable in any study dealing with translation product or process. However, it would be impossible to detect them and attempting to establish them would most likely be unethical. For more insight into methodological aspects of Internet-mediated research see Mellinger (2015).

To sum up, although studies researching professional translators in their normal place of work have started to gain momentum, to the best of my knowledge, freelance professional translators (apart from in Nord’s pre-Internet study) have not been studied before in their normal place of work. The present study offers a contribution to this area of research and attempts to fill this particular gap.

2.5. From academia to industry

The present study’s focus on human translation-oriented research activities has an additional, industry focus as, by trying to understand how professional translators interact with online resources, it attempts to provide a useful insight for technology developers who are interested in solutions aimed at reducing translators’ research time and improving their experience of using online resources. This focus is driven by the realisation that translation is no longer a ‘niche’ activity associated with literary works, but a multi-billion dollar business that is at the centre of the globalised economy, with freelance professional translators forming a large part of this industry. The translation industry in itself, as Mesa-Lao observes, “is going through a societal and technological change in its evolution. In less than ten years, the industry is considering new tools, workflows and solutions to service a steadily growing market” (Mesa-Lao, 2014:99).

However, whilst translators working in in-house settings at institutions, translation agencies or LSPs are naturally exposed to the new tools and workflows, freelance translators are very much left to their own devices when it comes to keeping up with developments in technology. Therefore, studying the way freelance professional translators use technology can provide a useful insight for technology developers as they form a community that is rather detached from the corporate or organisational environment (see e.g. Karamanis et al., 2011) and their practices are therefore less visible.

The shifting direction of research from purely academic foci to industry ones is very much aimed at studying the behaviour of translators in order to uncover potential links between this behaviour (including how they use tools and resources), the efficiency/productivity of the translation process (e.g. Moran et al., 2014) and the quality of the outcome (see e.g. Daems et al., 2015). Angelone et al., for example, note that
these new directions in research are often related to the examination of post-editing practices which, in turn, leads to an increased collaboration between machine translation and translation process researchers “with the common goal of enhancing productivity” (2016:50).

However, many of the developments are happening in the area of interactions with texts, i.e. during the translation activities (see e.g. Karamanis et al. 2011 for an overview of these developments), rather than in the area of interactions with resources, i.e. during the research activities. Admittedly, the technologies assisting with the interactions with texts focus very much on the leveraging of previously translated material so that there is less need for interactions with resources, but, as research shows (Daems et al., 2015; Zapata, 2015), translators still consult online resources even when using technologies such as machine translation. Although, to the best of my knowledge, no systematic studies empirically measuring the use of external resources with and without the use of translation memory have been carried out, the self-declared data from the Global Survey of the present study suggests that a comparable amount of time is spent on research activities regardless of whether or not the translators used translation memory or terminology management systems33 (see section 5.2.1.). This could be because the more content that is translated using TM and MT, the higher the potential for translation problems, thus the translation to research ratio might not change dramatically.

Therefore, with a view to streamlining research activities to increase productivity, researchers and practitioners alike have called for a tighter integration of online resources within the translation environment (Zapata, 2015; Zetzsche, 2015a & 2015b). Although search tools such as IntelliWebSearch34 and various plug-ins for CAT Tools35 are being developed, nevertheless, compared to the efforts in TM and MT development, very little large-scale integration of online resources has been on the agenda from software developers. Researchers, however, have made good attempts at trying to capture some aspects of resource use that could be used to develop better technology. Studies, especially in recent years, have also become oriented towards the development of new tools and resources or improvement of existing ones (e.g. Désilets et al. 2009; White and Matteson 2007; Massey & Ehrensberger-Dow 2010; Durán Muñoz 2012; Valli, 2012) in order to keep up with the demands of the market for ever faster translations in ever growing volumes. Nonetheless, more research is needed to produce findings that would be convincing enough for developers to respond to the needs that have already been expressed by researchers and practitioners for a tighter integration of online resources into translation environments. The present study also has an industry focus as, by examining human translation-oriented research activities, it is concerned with how these research activities can be streamlined in the future in a personalised way.

2.6. Other relevant research areas

The present study also draws on other research trajectories such as human translation styles, continuous development of skills associated with technology, i.e. continuous professional development (CPD) or the growing interdisciplinarity of translation process research, all of which will be briefly discussed below.

---

33 Similar percentages were observed for those participants using these technologies mostly/frequently and those not using them at all or using them rarely.
34 http://www.intelliwebsearch.com/
35 http://www.translationzone.com/openexchange/
Human translation-related styles

The importance of systematic research into individual differences in TPR has been recently highlighted by researchers, e.g. O’Brien (2013) in the context of translation or Timarová (2010) in the context of interpreting. Saldanha & O’Brien rightly point out that the individual differences relating to an individual’s personality, experience and attitudes can be observed in translation process research regardless of the degree of control over the experiment (Saldanha & O’Brien, 2014:147). And yet few studies examined individual differences (see e.g. Hansen, 2006 or more recently, Risku, 2014) and translation process-oriented research has mostly concentrated on looking for common tendencies and universal principles and has generated little research that explores and systematises differences rather than commonalities. This has resulted in a lack of insights and methodologies with which to study individual aspects of a translator’s performance and, as O’Brien pertinently concludes, many “intriguing questions that could be posed [...] have not yet been investigated” (O’Brien, 2013).

As mentioned in section 1.3., translation and research activities are perceived in this study from the observer’s point of view as two separate (although inevitably intertwined) activities within the translation process which involve different skills - translation and information-seeking respectively. It can be argued that this distinction was implicitly drawn by Carl and his colleagues (Carl et al. 2011; Dragsted & Carl, 2013) in their study of translators’ activity data carried out to “uncover characteristics of human translation processes” (Carl et al., 2011) in which, to the detriment of ecological validity as the authors themselves admit, the participants “did not have Internet access and were not allowed to use dictionaries or other similar support” (Dragsted & Carl, 2013:138). In what follows, a brief overview of studies particularly concerned with discerning patterns of individual behavioural characteristics in relation to translation and research activities will be discussed.

Human translation styles have been studied from two perspectives: the product perspective seen in earlier studies associated with translation criticism and the process perspective adopted more recently. The product perspective has so far been underpinned by a framework borrowed from psychology, whereas the process perspective of human translation styles has been borrowed mostly from information behaviour.

With regard to the former, Hubscher-Davidson points out that the discipline of translation studies has drawn on cognitive psychology in studying the mental processes of translators and interpreters, whilst the psychology of personality has been “largely overlooked” (Hubscher-Davidson, 2009:176). One of the few scholars who has commented on personality in translation is Reiss (1971/2014) who considered individual differences in the context of translation criticism. Based on Spranger’s six forms of personality (Spranger, 1920 in: Reiss, 1971/2014:110), Reiss concludes that some personality types are more suited to certain types of translation than others. For example, the theoretical type, thought to be dispassionate, objective and rational, is, in her view, better suited to translating technical and philosophical work than the aesthetic type, associated with a flair for artistic and aesthetic forms of expression. Further still, Reiss argues that some types, such as the aggressive type, regarded as egocentric, self-assertive, vital and vigorous, are not suited to translation at all based on their inability to empathise with the author’s original work (1971/2014:110). However, it should be noted that no empirical evidence is given to support these

36 Looking at individual differences rather than searching for universal principles within the groups being studied has so far been a strong theme in psychology and learning (in the context of translation in foreign language learning).
statements, made almost forty years ago. Nonetheless, the important aspect of Reiss’ approach relevant to the present study is the recognition that personality might predispose people to do things in a certain way. Furthermore, the personality types are only intended to point towards mind-sets or tendencies dominant in a translator’s personality (ibid.), which is a point pertinent to the present study in relation to translator research behaviour.

A more recent study examining the intersection of translation and personality was carried out by Hubscher-Davidson (2009). It was designed to gain an understanding of the influence of personality traits on the quality of translations. A total of twenty students were given a personality test (Myers-Briggs Type Indicator) and then asked to translate a literary text whilst verbalising their thoughts. Hubscher-Davidson found some correlations between the personality traits and the quality of translated texts. For example, “the intuitors amongst the students […] generally outperformed their sensing counterparts” (2009:184). The Intuitive/Sensing subscale37 was found to be the most promising for Hubscher-Davidson’s and for other similar investigations, for example in relation to students’ writing ability or to the general performance of engineering students (Hubscher-Davidson, 2009:185). However, it has to be taken into account that the students translated a literary text, which could have had an impact on the quality, since “translating literary texts requires a certain amount of imagination, and that taking a ‘practical’ approach would not readily result in a successful translation” (Hubscher-Davidson, 2009:186). Recently, a study which considers personality and its impact on translation and translators competence has shown that personality, although “not decisive in successful translation performance”, does influence translation process and product (Lehka-Paul & Whyatt, 2016)).

Moving away from psychology and the translated product and focussing more on the translation process, Asadi & Séguinot’s study looked for experience-related or text-specific strategies and the patterns exhibited by professional translators in a naturalistic environment. Based on their observations, they proposed two cognitive production styles: prospective thinking and translating on-screen. Translators who followed the prospective thinking pattern were “reading large segments of the text and making most of their translation decisions and changes mentally before typing” (Asadi & Séguinot, 2005:526) whereas the ones translating on-screen “spent less time planning, tackled shorter text segments and took advantage of the word processing environment to facilitate the process” (Asadi & Séguinot, 2005:530). Other shortcuts and strategies related to word processing and global strategies (drafting, drafting and checking, and revision) were also discussed. In contrast to the present study, Asadi & Séguinot were looking for universal strategies amongst the group, but apart from the two production styles, found that “each translator’s process is a unique combination of cognitive style, translating experience, technical skills and world knowledge, which cannot be fit into static categories we had hoped to find” (2005:539). However, the importance of this study for the present one lies in the recognition of the fact that “translation styles are not clear cut […] rather, it seemed possible to discern patterns38 and then deal with the groupings of patterns” (Asadi & Séguinot, 2005:525). The present study, although looking solely at translator research

---

37 Intuitors are those who focus on the meaning of life experiences and the relationships between them through imagining possibilities and sensing patterns and the sensors are those who focus on the more immediate and practical facts of life by noticing details (Pretz & Totz, 2007:1248).

38 Patterns are defined by Asadi & Séguinot (2005:525) as “a particular shortcut or strategy used by three or more translators”.

43
activities, not translation activities as such, adopts this approach, and as will be seen later, based on a substantial number of patterns, groupings become more viable.

In another process-oriented study, focussing on directionality in translation, Pavlović found that novice translators exhibited a particular combination of actions/interactions\(^{39}\) characteristic of their style of working, regardless of direction of translation (Pavlović, 2007:184). In Pavlović’s study each of the four groups who were working collaboratively on a task “had their own particular blend of actions/interactions that they employed - strategically or routinely - to complete the tasks” (ibid.). This observation gains even more importance in the context of the present study given the fact that the groups were not random, but self-selected. Each group comprised three students who had previously worked together in the classroom environment (Pavlović, 2007:70). It is therefore possible that the groups exhibited the patterns of the dominant person within the group or, indeed, the groups could have been formed on the basis of similar working styles. However, these are hypothetical assumptions not investigated by Pavlović.

Moving further into the translation process, researchers at the Copenhagen Business School (Carl et al. 2011:np; Dragsted & Carl, 2013) studied the translation behaviour of twelve professional translators and twelve MA translation students under strict laboratory conditions to look for patterns of behaviour during the translation process which excluded any interactions with external resources\(^{40}\). Using keylogging and eye-tracking techniques, they examined the unfolding translation process, seen as a process comprising an orientation phase (initial text planning phase), drafting phase (the phase in which the actual translation is drafted) and revision phase (the phase in which the translator reviews the translation and refines choices).

Based on their observations, Carl et al. (2011) proposed a taxonomy of human translation styles, in which translators were classified according to their observed behaviour in each of the phases. Thus they differentiated between:

**Initial orientation:** functions as an initial text planning phase. [...]:
- *systematic initial orientation*: the translator systematically reads through the ST before translation.
- *skimming*: the translator skims the ST rapidly before translation
- *quick planning*: the translator reads the first couple of words or sentences, and then presses the first key.
- *head start*: the translator starts translating right away

**Translation drafting:** the phase in which the actual translation is drafted. [...] 
- *large-context planning*: the translator reads text sequences, sometimes whole sentences, far ahead in the source text.
- *small-context planning*: the translator frequently fixates the word being typed or a couple of words, but rarely a whole sentence ahead.

**Revision:** this phase serves to review the text and refine translation choices. [...] 
- *online revision*: the translator revises the text during the drafting phase
- *end revision*: the translator spends 20 per cent or more of his/her time on end revision
- *constant revision*: the translator spends more than 20 per cent of translation time on end revision, but at the same time makes a large number (above average) of online revisions.

(Carl et al., 2011)

\(^{39}\) Pavlović defines actions/interactions as “everything that the subjects did in the course of their translation processes in order to transform the source text into a target text according to the assignment” (Pavlović, 2007:184).

\(^{40}\) As already noted, the participants in Carl et al.’s study (2011) had no access to the Internet and their usual tools and resources, and performed the task in an unusual situation, working with unfamiliar software. Although the authors believe that this setting did not influence the participants’ reading behaviour in any significant way, they do not address their position with regard to whether they believe it could have affected translators’ translation behaviour, the very object of the study.
Based on further analysis, Dragsted & Carl (2013:149) distinguished between two ‘styles’, i.e. a ‘local’ and a ‘global’ focus. The local focus is characterised by a very minimal or non-existent orientation phase, small/narrow-context planning in the drafting phase (i.e. focus on one or two words at a time) and online revision (i.e. revising alongside drafting), whereas the global focus is associated with a more systematic orientation, large/broad-context planning in the drafting phase (i.e. focussing on larger units) and end revision.

An important aspect of the above study lies in the comparison of the results made by the authors, with various writing style classifications such as those proposed by Bridwell-Bowles et al. (1987) (‘Mozartians’ and ‘Beethovians’) and Van Waes & Schellens (2003) (‘initial planners’, ‘first draft writers’, ‘second draft writers’, ‘non-stop writers’, or ‘average writers’) or Chandler (1993) (‘architect’, ‘bricklayer’, ‘oil painter’, ‘watercolourist’) to which classification Wyllie (1993) added ‘sketchers’. The main dimension of enquiry seems to correspond to the planning - discovering spectrum, where the ‘Mozartians’ are extensive planners who formulate and revise their texts sentence by sentence, whereas ‘Beethovians’ write a first draft of their text rather quickly with minimal revision, postponing the main revision until a later stage. The other categorisations of writing styles can be more or less placed on the planning-discovering spectrum, with different emphases put on the planning, drafting and revision phases. Despite noting some differences between writing and translation, Dragsted & Carl believe that “features characteristic of writing can be observed in translation as well” (Dragsted & Carl, 2013:137). Indeed, their findings point to the fact that the two main strategies employed by translators, the local and the global focus, can be linked to some of the categories found in writing styles. For example, the translators exhibiting the ‘local’ style were likened to Van Waes & Schellens’ (2003) ‘non-stop writers’ and ‘first draft writers’ or Chandler’s (1993) ‘watercolourists’ or ‘bricklayers’. The ones characterised by the ‘global’ style had more in common with Van Waes & Schellens’ (2003) ‘second draft writers’ and Chandler’s (1993) ‘oil painters’. With regard to the above-mentioned writing styles, Pym (2009) carried out a classroom experiment designed to raise awareness of translator styles based on Chandler’s classification (1993). However, whilst the students could relate to some of the styles proposed by and the experiment was a good opportunity for some “revealing self-criticism”, Pym struggled to relate the actual screen recording data to the corresponding styles (Pym, 2009:146). This could be due to the insufficient data and too shallow an approach to the analysis.

The taxonomy of translator styles provided an inspiration for the present study and a mapping of research styles on these translation styles was planned. However, in some cases not all of the phases of the translation process could be observed in the screen recording data of the Main Study (especially the orientation phase and in two cases final revision) and, when asked, some participants admitted they had a quick look at the text beforehand or would have spent a bit more time on revision. Therefore, only first steps were made towards exploring how research styles can be mapped onto translation styles (see section 7.2.6.).

---

41 In translation studies, Chandler’s classification is used by Mossop (2007).
42 Pym attributes these styles to Mossop.
43 None of them identified with ‘bricklayer’.
Continuous professional development (CPD)

With regard to the development of skills, it has to be said that so far no research has explicitly been carried out to address the professional development of professional translators, rather than the acquisition and development of competences of student translators. According to Pym (2009:136), there are several applications of empirical research into translators’ competence: longitudinal studies that aim to increase our understanding of how we acquire translation competence, comparison of student /novice translators with professionals or direct application of research methods and findings (such as screen recordings) to carry out in-class experiments as a means of learning about research and how it applies to real settings (Pym, 2009). However, there is no mention of professional translators who are already on their career path, but still require new insights for self-discovery, new skills and a better understanding of the processes involved in their everyday work, nor is there any mention of further development of competences beyond the classroom, in professional settings. And yet it is widely accepted that acquisition of expertise is a life-long process requiring constant updating of knowledge and skills. It also goes without saying that the needs of professional translators with regard to the development of various technology-related competences are different from the needs of students and therefore the design and goals of such studies should be different. Massey & Ehrensberger-Dow (2011a) do emphasise the importance of employing process research methods in translation practice to stimulate “reflection on decisions made and actions taken” and to “heighten[s] the awareness of key procedural aspects of translation practice and expertise” (2011a:7), but no explicit efforts have been made to focus on established professionals requiring further development, despite the fact that shortcomings in their technology-related competences have been empirically established (Massey & Ehrensberger-Dow, 2010:138).

As new technology-related competences will be constantly required of professional translators in future, research should embrace those needs in the ever faster changing professional and technological landscape. For example, Austermühl (2013) proposes that in view of the fast expanding battery of technology-based tools for translators, the focus in translator training should be shifted to revision skills (to do with post-editing MT) and documentation and research skills (to do with translation-related research activities). But empirical evidence is needed to explore these areas in professional settings in order to know not just how to teach these skills, but also how to integrate them into the already established environment, set of skills, beliefs and expectations characteristic of a professional translator (see e.g. Karamanis et al., 2011:11-12). It is hoped that the results of the present study will provide motivation to established professional translators for a much-needed self-reflection on how they make use of the wealth of online resources available to them as a means of improving their own performance.

Interdisciplinarity

With regard to interdisciplinarity, translation studies has always borrowed ideas from other disciplines such as literary studies, linguistics or philosophy. In the area of translation process research (TPR), research has also been crossing the boundaries between academic disciplines. From linguistics and psycholinguistics, psychology, neuroscience and cognitive science through writing and reading to language technology, human-computer-interaction and computational linguistics, researchers have tried to gain a “greater understanding of translation as an expert task” (O’Brien, 2013:1). Echoing Göpferich (2011), O’Brien (2013) rightly concludes that currently, translation process research can be seen as doing more ‘importing’ than ‘exporting’ of ideas and that a more symbiotic level of cooperation, reciprocal interdisciplinarity, or
Transdisciplinarity is the next step for translation studies. It has to be noted at this point that some criticism has been directed at scholars embarking on interdisciplinary research “without the corresponding background” (Hubscher-Davidson, 2009:176). An overview of the areas that currently influence translation process research is provided by O’Brien (2013); however, the influence of developments in language technology having already been extensively discussed in earlier sections of this chapter, it is the interface between translation studies and the larger, socio-cognitive context of ergonomics as well as translator profiling, mentioned by O’Brien (2013:8) in the context of psychology, that is of particular interest for the present research.

The interface between translation studies and ergonomics in the context of embodied cognition has been explored by Ehrensberger-Dow & Massey (2014) and Risku (2010). Ehrensberger-Dow & Massey (2014) have looked at how ergonomics can affect the efficiency of the translation process, the quality of the translated product or even the professional identity of the translator. The notion of ergonomics is extended beyond its common association with the physical dimension of a system to include organisational structures and, most importantly, cognitive aspects. It could be said that the present research fits into this trajectory by providing yet another aspect of translators’ ‘embodied cognition’, to do with how translators differ in their research practices, i.e. their resource and research style. This links to the second trajectory mentioned by O’Brien (2013:8), i.e. translator profiling based on their personality. Although not based on psychometrics but on online information behaviour, the present research offers a translation process-oriented profiling of translators which includes some aspects of personality.

To summarise, the present research has been shaped by the studies in translation process research and information behaviour. From the process-oriented studies, of greatest relevance were those that investigated translators’ use of online resources and their web searching behaviour as well as those examining professional translators as opposed to students and carried out in a natural environment. From the information behaviour studies, a much needed inspiration was drawn from those that provided various classifications of online information behaviours and considered the many possible lines of investigation into these behaviours.

2.7. Research questions

The present research responds to the many shifts and developments happening in the areas of translation technology and the associated global trends and attempts to contribute to the least researched areas within those developments. Firstly, it responds to the growing interest in researching translation technology, and takes a particular interest in online resources, seen as part of this technology. Secondly, it seeks to examine the use of these resources by freelance professional translators working in their natural environment rather than students of translation taking part in quasi-experimental studies in labs. It does so from the online information behaviour point of view, by looking at what they do, rather than from the cognitive perspective of what happens in their heads. In contrast to previous studies characterised by individual observations, the present research focusses on identifying patterns in an attempt to derive a consolidated view of possible types of translation-oriented research behaviour. Furthermore, the motivations of the present research are manifold and, apart from its pedagogical aims, it also has objectives associated with furthering professional development and technology development. Last, but not least, the present research responds to the under-researched area of individual differences in translation studies by looking at how translators differ in what they do rather than trying to find some universal...
principles. Based on the above review of the emerging trends and related literature, the present research aims to address the following research questions:

RQ1: What are the types of online resources available to professional translators and to what extent are they used by them?

RQ2: What proportion of their total translating time do professional translators spend on translation-oriented research activities? Does the actual time match the numbers they self-declare? How are the research activities distributed within the translation task?

RQ3: What patterns of behaviour (styles) can be identified with regard to the use of online translation resources by professional translators?

Having examined the relevant literature and formulated the research questions, the following chapter will explain the methodology that was developed to answer these questions. The chosen approach is a mixed methods approach and the above research questions are answered using a combination of two samples, one being a large sample of translation professionals that took part in the Global Survey and the other being a small group of freelance professional translators. Whilst RQs 1 and 2 are answered using both samples, RQ3 is based only on the small sample of freelancers.
3. Methodology

Having discussed the motivations for this study in Chapter 1 and the explanation of how it fits with previous work in the area of translation process research in Chapter 2, the present Chapter is concerned with presenting a methodology developed for this study to research how professional translators interact with online resources during the translation process.

In order to provide the first point of reference in relation to the design of this project (further elaborated on in section 3.3.), Figure 3.1 below represents its main components (Global Survey and the Main Study) with two components (Self-study and Pilot study) feeding into the Main Study, together with an indication of the two types of data analysis employed throughout the study.

![Figure 3.1 THE STUDY COMPONENTS AND ANALYTICAL APPROACH](image)

3.1. Research paradigm

The present study’s goal of gaining a better understanding of how professional translators carry out their translation-oriented research activities by means of online resources positions it within the process-oriented research tradition. The assessment of the translation product or its quality and cognitive considerations during the translation process such as translators’ thought processes, motivations or decisions are beyond the scope of this research. The predominant focus of this study is on the ‘behavioural’ aspects of translators’ research activities, i.e. on the observation and description of what professional translators do during their interactions with online resources, with the aim to provide some explanations as to why they might differ in the ways they conduct research for translation purposes. As such, the present study lends itself to an empirical approach based on various kinds of elicited data such as questionnaires, and direct observations by means of screen recordings accompanied by audio commentary.

The methods of data elicitation have been chosen based on the broader goals of the research design that were set in relation to the research gap identified in Chapter 2. The first goal is to study experienced, freelance professional translators, as opposed to novice translators or students, which is often the case in process-oriented research due to the convenience of obtaining subjects at universities and setting up the study. The second goal relates to ecological validity, i.e. studying experienced, freelance professional translators in their normal place of work as opposed to laboratory settings as elaborated on in section 2.3.
The above-mentioned goals lent themselves to the use of remote methods of data collection to enable freelance translators to perform the task in their natural work environment without the researcher being present during the task. However, they also made it necessary for the researcher to exercise a degree of control in setting up the Main Study - for example by providing a source text - as a completely naturalistic approach was not possible in the context of the research questions. This restricted the ecological validity of the study. Nonetheless, the least intrusive methods of data collection were chosen for the observational part of the Main Study in order for the study to resemble the natural process of translation as closely as possible (see section 3.2.2. for details).

Furthermore, the nature of the research questions invited the use of two techniques of data interpretation - numerical (represented in mathematical terms) and narrative (represented as common themes or categories). In other words, the study required a mixed-method approach, one that incorporates quantitative and qualitative approaches. As such, the present study lends itself to the pragmatic positon, a new paradigm recently developed in the philosophy of science. Within this ‘knowledge claim’ as Creswell (2003) puts it, “data collection and analysis methods are chosen as the most likely to provide insights into the question with no philosophical loyalty to any alternative paradigm” (Mackenzie & Knipe, 2006:3). Such a pragmatic position allows the researcher to overcome the quantitative-qualitative dichotomy associated with positivist and interpretivist philosophy respectively and is often seen as providing a framework for mixed-methods research (Tashakkori & Teddlie, 1998). This position allows research methods to be drawn from the social sciences as well from the natural sciences. With the research problem being central and the methods drawn “liberally from both quantitative and qualitative assumptions” that best meet the researcher’s needs and purposes (Creswell, 2003:11-12), pragmatist researchers look at the ‘what’ and ‘how’ which allows them to combine the objective and the subjective data in one approach. In the case of this study, self-declared and observed data is subject to triangulation, thus providing an aspect of validity to this research.

The pragmatic position enables the inductive reasoning approach to be adopted. This approach, based on observation, clustering, organising, the emergence and refinement of categories and the recognition of patterns, “open[s] up new paradigms generating new theoretical propositions” (Dancette, 2000:130), in the case of this study, in the form of taxonomies. Such inductive reasoning is associated with grounded theory, the basic tenets of which underlie the present research. These tenets, proposed by the founders of grounded theory, sociologists Glaser and Strauss (1967), and later adopted by many other disciplines, propound the discovery of a theory through a systematic research grounded in observation. In other words, rather than deducing (testing) a theory from the a priori assumptions, grounded theory strives to induce (generate) a theory a posteri or, from systematic research by means of comparative analysis. During this analysis “one generates conceptual categories or their properties from evidence; then the evidence from which the category emerged is used to illustrate the concept “ (Glaser & Strauss, 1967:24). At the heart of the grounded theory is coding, further elaborated on in section 3.2.3., by means of which conceptual categories, their properties and the relationships between them are established, thus facilitating the generating of a theory from the data. In the context of the above propositions, the next section will be concerned with the methods and instruments of data collection and analysis in greater detail.
3.2. Methods and instruments of data collection and analysis

In this section, firstly, the evolution of the methods of data collection and analysis in translation process research will be discussed within the context of the research design of the present study. Secondly, the methods of data collection used in this study will be discussed in the context of Krings’ (2005) classification\textsuperscript{44}, followed by an overview of what methods were used to analyse data in this study. Next, the instruments of data collection and the instruments of data analysis used in the present study will be presented, followed by the use of materials, i.e. source texts used for the translation task in the Self-study, the Pilot Study and the Main Study.

3.2.1. The evolution of methods of data collection in TPR

The methods of data collection and analysis in TPR, and, consequently, the instruments that facilitate this collection and analysis, have been evolving alongside technology developments, especially in Information and web technology. These changes are reflected in the evolution of research designs and the change of foci of the studies undertaken.

As far as the methods of data collection are concerned, before the widespread use of technology in the process of data collection, the translation process was investigated by means of ‘offline’ methods such as questionnaires, interviews or focus groups or ‘online’ introspection methods such as think aloud or talk aloud protocols. However, these methods on their own only provided “‘soft’, [i.e.] qualitative and subjective data” (Göpferich & Jääskeläinen, 2009:172). Observation of the actual behaviour could only be afforded by observing translators in person, by taking notes and later by video recording. With translation activities moving from ‘pen & paper’ to the screen and with resources moving from paper to online, the methods of recording these activities for research purposes have changed too. As Lauffer observes, “translation has changed considerably over the past few years, especially with the wave of new technology. Therefore, the way we observe translation must also evolve” (Lauffer, 2002:70). With the advent of computer-based methods which were capable of delivering “‘hard’, quantitative and objective data” (Göpferich & Jääskeläinen, 2009:172) a more scientific approach was welcomed and methods such as Internet logging, keylogging, screen recording or eye-tracking became the main methods of data collection in process-oriented research as they allowed researchers to capture the various aspects of the translation process as it unfolded (O’Brien, 2011). Methods from neuroscience, such as electroencephalography (EEG) and neuro-imaging techniques, such as functional magnetic resonance imaging (fMRT) and positron emission tomography (PET) have also been applied (see e.g. Neubauer & Kanape, 2009). However, while methods from computer science such as Internet logging or screen capture have proven extremely useful in capturing aspects of the translation process that so far have been extremely difficult or impossible to access, the neuroscience methods have proven to be less productive for investigating the translation process, as they mainly provide cues as to which areas of the brain are involved in any particular activities (Göpferich & Jääskeläinen, 2009:170).

\textsuperscript{44}In this study, methods referred to by Krings as methods of data analysis (questionnaires, TAPS etc.) are considered to be methods of data collection. Methods of data analysis in this study refer to methods such as descriptive statistics, coding or triangulation.
The classification below, developed by Krings (2005:348, translated by Dam-Jensen & Heine, 2009:3) and complemented with the more recent methods (added in bold), provides a systematic overview of the current methods of data collection used in translation process research.

However, it is not only the methods of data collection that have evolved. The methods of data analysis have also undergone a transformation. Traditionally, methods of data analysis have been divided into qualitative and quantitative, qualitative methods being associated with research carried out in arts faculties, social sciences and education and quantitative research being broadly associated with research in the natural sciences. During the 70s and 80s they were thought to be “fundamentally incompatible […] [which] created a dilemma for researchers who used methods of both qualitative and quantitative orientation in their studies” (Symonds & Gorard, 2010:3).

As mentioned above, with the advent of computer-based methods of data collection and thanks to the influence from research in psychology, it became possible to introduce a more objective, quantitative element of analysis into process-oriented translation research, which had hitherto been dominated by qualitative analysis based on analysis of the product and TAPs. Researchers have started to combine these two approaches and in order to provide a methodological justification for the use of multiple methods, the notion of triangulation was introduced as an approach to dealing with different data types.

Today, the traditional dichotomy between qualitative and quantitative methods of data collection has been challenged and in many disciplines the mixed-methods approach seems to be gaining momentum. As Onwuegbuzie & Teddlie (2003) observe, a new movement called the ‘third paradigm’ which explicitly combines the qualitative and quantitative approaches has emerged. Johnson & Onwuegbuzie (2004:17)
define the mixed-methods approach as “the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study”. For an alternative view see Symonds & Gorard (2010)\textsuperscript{45}.

Although it could be said that the mixed-methods approach to research has gained in popularity, it is known that this approach is not entirely new. In the past, many projects have naturally combined or integrated methods by coding and counting qualitative data and by the interpretation and explanation of quantitative data, without explicitly stating that this was a special approach to carrying out research (Kuckartz, 2010:3; Plowright, 2011:2-3). What is new, however, is the technology that supports mixed-methods research, which had not been available before and which is capable of not only collecting but also analysing vast amounts of data, far exceeding the scope of earlier research. The range of types of data that can be directly handled by data analysis software has dramatically increased in recent years. Now, many types of multi-media data (e.g. still and moving images, web pages or Google Earth records) as well as any form of textual data, including pdf files, can be directly handled by most data analysis software applications. One example of such technology is MAXQDA\textsuperscript{46} software, which is used in the present study.

To conclude, not only the methods of data collection but also the methods of data analysis in translation research have changed over time. The qualitative-quantitative dichotomy has been weakened and a mixed-methods approach is gaining in popularity, with researchers starting to use technology not only to collect data but also to analyse it. The abundance of analytical tools capable of sophisticated investigation of any kind of data, be it textual or multimedia, helps to broaden the scope of research and deepen its analysis. So, although there is some criticism regarding the legitimacy of the mixed-methods approach as a separate paradigm, due to the complexity of the phenomenon of the online behaviour of professional translators studied in the present research and the variety of insights sought with regard to the patterns of this behaviour, it is considered that this approach is the most suitable for answering the research questions posed by the present study.

3.2.2. Methods of data collection used in the present study

A comprehensive overview of methods used in data collection in translation process research has been provided by others (see for example Dancette, 2000; Lauffer, 2002; Hansen, 2006, 2008; Dam-Jensen & Heine, 2009; Saldanha & O’Brien, 2014). The methods that were chosen for the present study, shown in the unshaded area of Figure 3.3 will be discussed below.

\textsuperscript{45} Symonds & Gorard (2010) strongly criticise this new paradigm. They point out that the explicit need to mix qualitative and quantitative methods to overcome the limitations of both approaches actually reinforces the qualitative – quantitative dichotomy. Symonds & Gorard (2010:17) propose that all paradigms should be abolished and a new approach to research – “plain ‘research’ as craft” is born, in which all methods have a role and a place in the full research cycle.

\textsuperscript{46} http://www.maxqda.com/
Online questionnaires

In Krings’ classification (see Figure 3.2) questionnaires belong to the verbal-report data, collected offline, i.e. not concurrently with the translation process (Krings, 2005:348, in: Dam-Jensen & Heine, 2009:3). In the case of online questionnaires, they can be classified as ‘mechanical’ due to the data being collected through an Internet data collection tool. Questionnaires have been used in translation process research in numerous studies to study the actual use of online resources by students, novice translators and professional translators (Hirci, 2012, 2013; Massey & Ehrensberger-Dow, 2011b; Pinto & Sales, 2008; Sales & Pinto, 2011) or to gain insight into their perceived needs and expectations with regard to these resources (Durán-Muñoz, 2010, 2012).

In the present study, questionnaires were used in the Global Survey to ascertain the general habits and preferences of professional translators with regard to the use of online resources. Questionnaires were also used in the Main Study as a tool to gather background information about the participants as well as particular information about their self-declared use of resources, time spent on research activities, attitude towards technology etc. which will be needed for data triangulation with the observed data from the Main Study.

The main problem with collecting data through this method occurred at the design phase, especially in the Global Survey. A decision was required as to whether to force participants to answer all questions, thereby providing complete but possibly unreliable data or whether to give translators the freedom to skip questions, thereby avoiding the risk of obtaining potentially untruthful data, though almost certainly yielding incomplete data. Online Survey Design Guide (online) advises that “forcing respondents to answer questions should be used only on rare occasions because the respondent may become annoyed and give an arbitrary or deliberately false answer in order to proceed to the next screen or stop taking the survey altogether”. Therefore, discarding unfinished questionnaires was considered more advantageous than working with data that could have been potentially untruthful.
Screen recordings

Screen recording (also referred to in the literature as screencast or screen capture) is a ‘mechanical’ observational method of data collection and belongs to the online methods, i.e. those where data is produced at the time of the translation process (Krings, 2005:348, in: Dam-Jensen & Heine, 2009:3). It has been used in translation process research to study writing activities, revision processes or decision making, to name a few. The screen recording technique has also been used to study research activities during the human translation process (Lauffer, 2002; Asadi & Séguinot, 2005; Ehrensberger-Dow & Massey, 2008; Enríquez Raído, 2011; Ehrensberger-Dow, 2014; Mutta et al., 2014) and during the post-editing process (Zapata, 2015). For example, in her study of three professional translators, Lauffer used Camtasia software to see “how they consulted electronic dictionaries, databases, websites and other tools” (Lauffer, 2002:63). She found that the screen recording software “was most useful in following the search paths and helping understand how and why the translators searched for information” (Lauffer, 2002:69). In this study, screen recording was used precisely in the context outlined by Lauffer.

Screen recording was chosen as the preferred method to carry out an observation of the participants in action during the Pilot Study and during the Main Study in order to:
- provide observed data to complement the self-declared data to answer RQs 1 and 2 related to the time spent on research activities and the types of resources used during these activities and
- provide the main dataset to answer RQ3 related to the patterns of use of online resources.

The use of screen recordings in the quasi-naturalistic setting of this study was made possible by the fact that it has become technically feasible for participants to access web-based screen recording software directly from their browsers, without the need to install software on their computers. Other ‘mechanical’ observational methods such as eye-tracking or keylogging were ruled out as it would have been extremely challenging to install suitable software on the personal computers of the participants and no web-based solutions are to this day available to the best of my knowledge. As the participants were located in various parts of the world, direct observation by the researcher was also ruled out for practical reasons. In hindsight, a web-logging tool would have been beneficial for this study, as it would have enhanced the data by providing additional statistical information such as the exact time translators spent on each individual resource. However, the data calculated manually provided enough information for the purpose of this study.

Screen recording software records all on-screen activities so that the entire process of translator interaction with online resources can be post-examined in the context of the translation task. As with any other method of data collection, screen recording has advantages and disadvantages. However, it is generally considered one of the best methods of direct observation, especially owing to its invisibility and non-intrusiveness.

Originally developed for training and demonstration purposes and widely used in areas such as education, screen recording has become a very useful research tool used in translation process-oriented research. As Göpferich and Jääskeläinen observe, methods such as screen recordings is “particularly useful for analysing the research activities which form an integral part of translation processes, as they provide a detailed account of which electronic sources or websites the subjects are using during translation” (Göpferich &

47 https://www.techsmith.com/camtasia.html/
Jääskeläinen, 2009:173). An important attribute of screen capture software, especially in the context of this study, is the fact that after the initial set-up, it runs quietly in the background and, as Asadi & Séguinot (2005:523) point out, “does not affect the translator’s natural working environment, an important factor in maintaining the ecological validity of the data”. Despite its obvious advantages, however, using screen recording can still be considered as introducing “a foreign element into the translation process [...] [which] may consequently interfere with said process”, as Enríquez Raído (2011:132) observes. She adds that, as in the case of using video camera recordings, “subjects may feel uneasy being observed and hence may change their behaviour” (ibid.). The ‘observer effect’ applies to most research with human subjects thus making it almost impossible to carry out an ecologically valid study in which the subjects’ behaviour is at their most natural.

To counter this argument in relation to the use of screen recordings, Pavlović (2007) mentions a study from second language acquisition (reported in MacIntyre & Gardner, 1994) in which two focus groups reported similar levels of self-reported anxiety, despite the fact that a video camera was used only in one of the groups. “Although other studies have produced different results, this at least warns against taking for granted the anxiety-inducing effect of video recording”, Pavlović (2007:52) explains. Other drawbacks pointed out by researchers are the fact that accurate information on the attention units cannot be collected and the lack of logging facilities within screen capture software that could provide numerical data for the analysis (Sullivan & Lindgren, 2006). However, in future research these drawbacks can be overcome by combining screen recording with other methods of data collection such as keylogging, eye-tracking or any offline method, such as verbal reporting or TAP. Also, new types of user experience and market research software such as Morae could fill the gap. Another valid and important drawback pertinent to the use of screen recording as a method of data collection in this study relates to confidentiality issues. One of the participants mentioned that personal or confidential information could appear on the screen as pop-ups during various online activities such as dedicated e-groups, Skype and other social media notifications that appear on the screen in real time. Although this did not happen during the translation task in the Main Study, another participant who intended to use Skype as an ‘ask and expert’ tool, seemed to have changed her mind, very likely for this particular reason.

Nevertheless, despite the drawbacks of this method, most participants in this study generally responded positively to using the screen recording technique and reported that it did not make any difference to the way they worked: in the post-task questionnaire, participants reported that “the use of the software had no impact on my work. It did not interfere in any way” (T3), “it was unobtrusive” (T12) or “I don’t feel it interfered with my normal work” (T16). Only one out of the 16 participants (T9) explicitly said that “being recorded and knowing that somebody else is going to listen to it made the difference”. Some other translators said that they were “slightly affected” (T8) or reported that they felt awkward at the beginning but as the task progressed they forgot that their activities were being recorded (T11 and T13).

48 https://www.techsmith.com/morae.html
49 Translators taking part in this study were anonymised, so for example T3 stands for Translator3.
Think/talk aloud protocols (TAP)

This method of data collection is classified by Krings (2005:348, in Dam-Jensen & Heine, 2009:3) as an online, verbal-report method. Before the advent of computerised methods of data collection, talk aloud/think aloud protocols were the most widely used methods of data collection in process-oriented translation research (Dam-Jensen & Heine, 2009:10). Talk aloud and think aloud, although slightly different, can both be thought of as the concept underlying the verbalising of one’s thoughts and actions while performing a task and will both be referred to as TAP.

So far, TAP methods have been used mainly to gain access to the mental processes that take place during the translation task. In this research, however, the TAP method has a slightly different purpose. It was used as a supplementary method accompanying the screen recordings in the Main Study to gain additional information about the translation process during the translation task such as evaluating the types of translation problems, commenting on the level of satisfaction with the solution or commenting on the usefulness or functionality of the resources to answer RQ3. Therefore, in this study ‘audio commentary’ is also used interchangeably to refer to the same activity.

Dam-Jensen & Heine (2009:10) aptly enumerate the shortcomings of TAP studies, which they subsume under three headings: accessibility, incompleteness and interference. Accessibility, they explain, refers to the fact that it is impossible to access people’s minds and that, essentially, data gathered this way can “only be considered indirect data [...] [as one] cannot know whether verbalizations are in fact a reflection of what goes on in the mind” (ibid.). In the case of the current research, this shortcoming does not bear much relevance as the purpose of verbalisations is not to access cognitive processes, but to illustrate some of the aspects of the process from the behavioural perspective.

The second criticism of TAP - incompleteness - refers to elicited data which often appears to be fragmented and subject to the participant’s decisions as to what is being verbalised and what is not (Hansen, 2005:516). Studies show (Kiraly, 1995; Jääskeläinen, 2002; Hansen, 2005) that the deeper, subconscious processes are apparently not verbalised. Also, an important observation by Séguinot (1989) suggests that highly experienced professional translators do not necessarily produce many verbalisations when they perform their routine tasks (reported in Tirkkonen-Condit, 2002:6). This could be due to the fact that many of the translation professionals’ processes are automated with practice and repetition, and, therefore, cannot be verbalised (Asadi & Séguinot, 2005:522). Asadi and Séguinot also comment on Ericsson and Simon’s studies of 1980 and 1984 (2005:522) in which “subjects [...] [could] only verbalize thoughts in short-term memory”. Automated skills, they conclude, “may bypass this short-term memory” (reported in Asadi & Séguinot, 2005:522).

This criticism is a valid one, but one must bear in mind that most TAP studies rely on participants’ spontaneous verbalisations, i.e. no instructions are given as to what aspects of the process should be commented on and how. This could also be a reason for the incomplete data obtained from TAP, as participants might not be clear about what exactly they are required to comment on. In the case of the present research, participants received instructions to comment on particular aspects of the process. Nevertheless, this has not proven to be a remedy for the incompleteness of the information obtained as participants did not always remember to comment or did not comment on certain aspects at all, even with the instructions given.
The third criticism of TAP relates to interference, i.e. the fact that the process of verbalisation disrupts the natural translation process and that “it influences the sequence of thoughts” (Dam-Jensen & Heine, 2009:10). Some scholars, for example Toury (1995:235), believe this to be the case, arguing that writing and speaking are governed by two different cognitive processes and, therefore, could be causing interference in the translation process. Ehrensberger-Dow and Künzli (2010:121) also report that TAP can significantly slow down and change the translation process. However, others, such as Ericsson (2006) or (Sun, 2011) claim that verbalisation does not affect task performance. So, in conclusion, Dam-Jensen and Heine argue that it is not quite clear how TAP influences the process, if at all (Dam-Jensen & Heine, 2009:10-11). In the present study, the fact that participants were asked to reflect on their research activities could have had an impact on their research time. This was partly addressed by subtracting the time where comments were made with no observable research activities; however it was not possible to account for the possible slowing down of research activities accompanied by audio-commentary. To address this further, a question was asked in the post-task questionnaire to ascertain whether the TAP component was perceived as having detrimental effect on the process.

In the present study, TAP was conducted during the Self-study, the Pilot Study and then in the Main Study. In the Main Study, it was used concurrently with the screen recording, so both audio and video files are linked together. In the Pilot Study, it was revealed that both participants had a completely different experience with TAP, and hence conflicting opinions as to the extent of the impact of TAP on the translation process, corresponding to second criticism above. During the post-task e-mail questionnaires in the Pilot Study (see Appendix 1), both participants were asked: “Did you feel that the act of talking aloud during the translation task has had any impact on the process? If yes, please say in what way”. Below are their accounts:

Katarzyna: “It was rather interruptive and I had to remind myself to talk through what I am doing. I usually think quicker than talk and it also slowed me down a bit”.

Saskia: “I talk to myself while translating quite a lot anyway - definitely not as much as during this translation task, but I do not think this affected my translation process in any way”.

The Main Study participants’ reactions to TAP were also almost equally divided into two groups. A total of 44% of the participants reported that TAP affected their performance, whilst 56% said that it did not, with some saying that it was actually helpful.

For example, T1 reported that “Yes, talking out loud slows you on the process of translation plus many things you do “automatically” normally must be made consciously, and explained, which obviously takes away from the “flow” of translation. The sequential steps you normally follow in your brain are interrupted by explaining the process itself”. T5 found herself “not knowing what to say. It was difficult to me to put in words what my brain was processing”. On the other hand, T9 reported that “talking out loud itself didn’t have much impact on translation”. T15 also “did not feel it really had an impact on the process as I sometimes think aloud when I work anyway”.

However, one participant pointed out the fact that although she thinks she was not able to express her ideas clearly, she did not think this was due to the conditions of the experiment. Rather, she said, “I sometimes find it difficult to verbalise my thoughts in a clear and organised manner whatever language I use. I much prefer writing in general”. This could point to the fact that the successful elicitation of verbalisation is not only dependent on external conditions but could be due to a host of other reasons, for example, the preferred medium of expression. While it is clear from these accounts that TAP had a
different influence on the translation process for each participant, and there is a possibility that some individuals could be more predisposed to verbalisation than others (also see Lauffer, 2002:64-65), it is not clear whether the verbalisation process had any impact on the participant’s interaction with external resources.

What is clear, however, is that TAP can raise awareness of the processes translators engage in. Although 37% of the participants said after the task that TAP did not raise their awareness during the task, 63% declared that it actually did. For example, T1 said: “yes, I found that many of the things I do in my own translations are automatic and I was verbalising them as I was doing them, which gave me a heightened sense of awareness of what I was doing and why I was doing it”. T15 reports that TAP “actually did encourage me to think deeper about what I was doing whereas I would generally do all this quite mechanically, without second thoughts about what I am actually doing”. This could have possible implications for training as classroom translation practice rarely facilitates the use of voice during the class.

A couple of participants related the benefit of talking aloud to being able to verify the naturalness of their translations. For example, T3 said: “I find talking beneficial. Sometimes it is easier to judge if a translation sounds natural if you say it aloud” and T12 reported that “a couple of times I backtracked and modified my initial translations after speaking about them and feeling that they didn’t sound right”. This finding becomes relevant when it comes to the adoption of speech-to-text technologies. Translators who find it beneficial to see and listen to how their translations unfold will probably find these technologies a great asset in their productivity strategy.

As mentioned above, whilst these observations are relevant for assessing whether TAP is a promising method for studying translation activities during the translation process, they actually do not say anything about how TAP influenced the research activities, i.e. the interaction with online resources. In fact, the major disadvantage of using TAP in this study was to do with calculating the duration of translation and research activities as much ‘commenting’ happened between these two types of activities as well as overlapped them as translators talked through what they were doing. Therefore, the job of separating the entire translation task into translation, research and commenting activities was one of the most challenging tasks in this study. It could not have been performed automatically and this has to be taken into account in future studies where TAP is used concurrently with screen recordings and where statistical data on the duration of various activities is sought.

Post-task e-mail questionnaires

As an offline, retrospective verbal-report data collection method, post-task email questionnaires were used in this study in order to help answer RQ3 by achieving a greater depth of understanding of each individual participant’s case. It was also employed to evaluate the overall experience of this study in terms of ecological validity, as well as to provide an opportunity to ask any additional questions or fill in the gaps where incomplete data was presented.

Retrospection has a long-standing tradition in process-oriented studies. Krings (2005:348, in: Dam-Jensen & Heine, 2009:3) lists two main types of retrospection: retrospective comments and retrospective interviews/questionnaires. Hansen (2006) reports on the many types of retrospective comment methods, and a detailed overview of retrospection as a research method in studying translation and simultaneous interpreting process is provided by Englund Dimitrova & Tiselius (2009). Their conclusion that retrospection
cannot be used as a stand-alone method and is best combined with other methods supports the way post-task questionnaires were used in the present study, i.e. as a source of supplementary information.

Instead of conducting face-to-face, Skype or telephone interview, an e-mail questionnaire was chosen for the Pilot Study and as the final component of the Main Study due to its many benefits. Firstly, this method is extremely convenient for both the researcher and the participant, as there is no need to arrange a meeting or travel, and participants can respond in their own time. This also includes cost saving (time investment, travel costs) and avoiding having to undertake a risk assessment if entering a participant’s home. Other benefits include the fact that the time-consuming process of transcribing the interviews is avoided and the data can be analysed straight away. It has also been noted that although “the tidy nature of the e-mailed response does not stand comparison with the engaging candour of remarks made spontaneously within the context of a successful face-to-face interview” (Lowndes, 2005: 130), people often feel more comfortable being in full control over what is being recorded, by being able to edit their responses (ibid.). Also, this method bypasses the need for good interpersonal skills such as eye contact and active listening which are required of the successful interviewer (ibid.). All in all, the flexibility of this form of retrospective data collection and all the above-mentioned benefits have contributed to the choice of post-task e-mail questionnaires to complement the screen recordings in the Main Study of this research.

Although the use of the email questionnaire as a complementary research method worked well, it must be said that it has proven to have its drawbacks. For example, unlike in the interview, the researcher has no control over the way answers are provided and although it is possible to ask for clarification, there is no guarantee that the answers given in the follow-up emails will be satisfactory (see for example footnote 206 in section 7.1.2.5). The answers given in the interviews were often short and although they were to the point, more insights could have been obtained in an interview. Also, if a participant cannot recollect certain information, the researcher is not there to give a cue and therefore the data could still remain incomplete or the opportunity to obtain additional information missed. Therefore, it can be said that the convenience for the researcher and the participants afforded by email questionnaires comes at a cost of obtaining data that is less rich and potentially lacking depth and clarity that follow-up questions in a true interview technique would allow.

3.2.3. Methods of data analysis used in the present study

As mentioned in section 3.1., the methods of data analysis used in the present study incorporate both the quantitative and qualitative approaches in the form of descriptive statistics and coding respectively. Triangulation techniques are also used to facilitate validation of data by means of cross verification from two or more sources. These methods, shown in the unshaded area of Figure 3.4 below will be discussed.
Descriptive statistics

Descriptive statistics are used in all components of the present study to present quantitative data, mainly in the form of summaries and graphs using counts, ranges, percentages, mean values, weighted averages and standard deviation to illustrate the general tendencies of the sampled populations as well differences between individuals within the population.

Inferential statistics were not used as it was felt that the sample, especially the one in the Main Study (N=16), was too small to make generalisations with regard to the population of professional translators estimated to be 250,000 (TAUS, 2011:online). Moreover, for inferential statistics random samples work best, while in this study the sample was selected based on specific criteria, i.e. the length of experience and language direction, and, therefore, was not random. Finally, the focus of the present research on the relatively new territory of the online information behaviour of professional translators, suggested that a small group should be studied in depth as a first step and the identification of central tendencies, spread etc. be carried out first by means of descriptive statistics. It was also important to combine this with qualitative analyses in order to identify initial patterns and to paint a broader picture. An important purpose of this type of study is that it helps us to identify the relevant variables for further study. The use of inferential statistics requires a tighter control of the variables in order to make valid claims about the population based on a sample and, in a relatively new field of study where little is known about the nature of a process or the behaviour of the group, this would not have been possible without learning more about those variables.

Although inferential statistics were not used, a degree of external validity associated with their use was achieved by conducting this study as a language independent one in order to observe a broad spectrum of

---

50 Apart from calculating z-scores in some instances to make decisions about outliers.
resources being accessed in as many target languages as possible and through triangulation. For example in answering RQ2 regarding the time spent on research activities, the self-declared data from a population of 540 from the Global Survey was compared to the self-declared and observed data from the population of 16 from the Main Study sample.

**Coding**

Coding was used in the process of analysis of the screen recorded translation task in the Main Study as a means of constructing an analytical framework for answering the open research question No3: ‘What patterns of behaviour (styles) can be identified with regard to the use of online translation resources?’. As no assumptions or hypothesis were made as to what these patterns could be, the approach was to observe the actual behaviour of the participants and arrive at the patterns through the analytical process of categorising data to facilitate the analysis based on grounded theory. Coding enabled the researcher to carry out two types of analyses: to qualify the quantitative data and to quantify the qualitative data.

The first type of analysis, qualifying quantitative data, resulted in the separation of the time-related features of the screen recorded translation process with its audio commentary into three types of activities: translation, research and commenting. Furthermore, the time-related features of the research activities were categorised into research step, research episode and research session, which will now be defined and discussed together with other elements of the analytical framework of this study. The definitions are adopted from translation studies as well as from information retrieval, a branch of information science that concerns itself with web searching. These definitions are bound by the methodological implications of this study, i.e. they refer to what can be directly observed from the screen recordings and heard in the audio commentaries without taking into account participants’ cognitive activities, such as problem-solving or decision making, unless they were otherwise verbalised during the screen recording.

**Translation episode**

*Translation episode* refers to the time spent on an uninterrupted interaction with the text in the form of translation and is observed when the text window can be seen on the screen recording. A translation episode ends whenever the translator switches to a resource window.

**Figure 3.5 SCREENSHOTS OF TEXT AND RESOURCE WINDOWS**

Screen recording – screenshot of text window

Screen recording – screenshot of resource window
In some instances, a problem occurred when the translator had both a source text window and the resource window displayed simultaneously on the screen.

This made it difficult at times to tell exactly when the switching occurred, especially when ‘backtracking’ took place (see section 5.3.2.) and an eye-tracking technique would have been helpful to measure the ‘switching’ in this instance. However, in most cases, it was possible to measure the switches with reasonable accuracy due to the visibility of the mouse cursor.

Whenever it was possible to clearly distinguish the commentary from the text production or research activities, the interruptions caused by audio commentaries were ‘calculated out’ of the translation episodes (and research episodes for that matter). Whenever the commentary ran parallel with the translation or research episode, it was not calculated out. However, if a participant declared in the post-task questionnaire that the talking aloud had a considerable impact on the translation process (which was the case with four of the participants) this was noted in the qualitative analysis.

**Research need, research unit, research session, research episode and research step** (see Figure 3.7) are encompassed in a research action and will be defined below, starting from research step and working up towards research need.

![Figure 3.6 SIMULTANEOUS DISPLAY OF TEXT AND RESOURCE WINDOW](image1)

![Figure 3.7 THE STRUCTURE OF A RESEARCH ACTION](image2)
Research step

*Research step* occurs whenever a resource is accessed in response to a *research need* (defined below). ‘Accessing’ is defined by the act of ‘clicking’, whether directly from a taskbar, through a web browser or from within a particular resource by means of hyperlinks. Each instance of ‘clicking’ into another resource or a sub-section of a resource constitutes a ‘step’. For example, accessing Wikipedia is considered as step 1 and accessing a language version of Wikipedia is considered as step 2. Similarly, accessing Google counts as step 1 and accessing a page from Google results counts as step 2.

A problem occurred, especially with Google results, when translators consulted the ‘snippets’, i.e. the description text appearing below links in search results (highlighted in yellow in Figure 3.8 below) but the consultation did not involve going into the resource, i.e. ‘clicking’ into it. These instances of consultation were not accounted for as it was not always possible to determine from the audio commentary which Google ‘snippet’ was consulted. The eye-tracking technique could have been of assistance in this case.

![Google search results with snippets highlighted](http://www.wikipedia.org/)

**Figure 3.8 ‘SNIPPETS’ IN GOOGLE SEARCH ENGINE RESULTS**

Research episode

*Research episode* refers to a series of steps (as defined above) taken to address a particular *research need* and is tied to a particular lexical item in the source text, i.e. a *research unit*. This includes accessing new resources and any modifications to research queries and submitting new queries within the same resource.

Research episodes have a different ‘construction’ from translation episodes. While a translation episode is an uninterrupted act of engaging with the source text and producing target text and its end is communicated by an observed switch indicating that interaction with an external resource has commenced, a research episode does not necessarily end when a switch on the screen is observed. This is because multiple switches between resources are common during the research process and all of these switches to a different resource or a different part of the resources by means of hyperlinks or bookmarks form part of a single research episode. The individual switches within a research episode are called *research steps*. Therefore, while a translation episode is characterised by one continuous interaction, a research episode can be made up of many interactions with various resources.

---

51 http://www.wikipedia.org/
Research session
Research session refers to a series of research episodes referring to the same research unit, separated by more than one translation episode. For example, whenever a research unit is researched at different times during the translation process, such as during the orientation phase and then again during the drafting phase (Dragsted & Carl, 2013), these two research episodes form a research session. (Vanessa Enríquez Raído, 2014:127) refers to a research session as a “temporal series of online actions aimed at satisfying a specific information need”, but it is not clear whether it encompasses searches that were triggered by the same research need but happening at different times during the task. Therefore, a unit called research episode was introduced as the main unit of analysis.

Research unit (RU)
While research need refers to the kind of information required at a particular point in the source text (e.g. target language equivalent, spelling, capitalisation etc.), it is important to acknowledge that the needs are triggered by particular words or phrases in the source text. In this work, they are called research units. Therefore, research unit refers to a lexical item (a single word, a part of a word, or a chain of words that acts as a unit of meaning) in the source text which is being researched using external resources during the translation process, in response to a research need. The research unit can be a word (i.e. general language), a term (i.e. specialised language), a proper noun or a phraseology unit.

In the Main Study over 80% of research units were explicitly indicated by the participants, either by verbalisation or other means, such as highlighting the research unit in the text. The remaining 20% of research units that were not explicitly indicated by the participant were established on the basis of the research queries that were typed into online resources. Seven translators explicitly indicated all the research units which they researched. Given the relatively small percentage of the research units that were not explicitly indicated, all RUs were given equal status as if they were all explicitly indicated.

Research need
Research need in this study corresponds to Enríquez Raído’s search need, i.e. “the recognition of an information need as perceived within the context of translation problem solving” (Enríquez Raído, 2011:152). Research need refers to the nature of information required, such as understanding the concept, finding an equivalent, checking spelling etc., and implies that research in external resources follows the identification of the need. When the flow of translation activity is interrupted and the translator looks to external resources to solve a translation problem encountered in the source text, a switch from a text to a resource occurs. This switch normally indicates that there is a research need that is closely tied to the actual word or expression in the text called a research unit. This triggers the use of external resources and a research episode follows. The research design of the present study included provision for finding out more about translators’ research needs. For example participants were asked to say whether the problem they are trying to solve relates to unpacking (comprehension) procedure, which is the process of understanding of the source text concept as embedded in its source language and culture, or whether it relates to repacking (reformulation) procedure, which is the process of expressing the source text concept in the target language to fit the target language's linguistic, conceptual and cultural context. However, the

---

52 This is assuming that no solution was found in internal resources (i.e. translator’s memory).
53 In some cases this switch could not be directly observed - see definition of a ‘translation episode’ for details.
participants do not always provide this information in the audio commentaries. Therefore, due to incomplete data, this aspect could not be analysed.

The second type of analysis, the quantifying of the qualitative data, was based on counting the number of occurrences of each code assigned to a certain behaviour observed in the screen recordings. The coding procedure was as follows: the categories of analysis that emerged from the Self-study and the Pilot Study (discussed in more detail in sections 3.3.2. and 3.3.3.) were used as basic codes capturing the main emerging themes. Then sub-codes, which revealed more specific patterns within these themes, were identified in the screen recordings. For example: ‘search strategy’, identified in the Pilot Study, was used as a basic code and ‘top-down’ and ‘bottom-up’ as sub-codes. However, when it was observed that T9 had repetitively used a combination of the ‘top-down’ and ‘bottom-up’ strategy, a third sub-code, ‘hybrid’, was added, in keeping with the grounded theory approach. Other codes (and their sub-codes) corresponding to the categories previously emerging from the Self-study and the Pilot Study included: research sequence (parallel, sequential, meandering), research path (straight, circular, meandering), research direction (ST-oriented, TT-oriented), research depth (shallow searches, deep searches).

Any additional characteristic behaviours observed received a new code, for example, ‘squirrelling’, ‘drive for perfection’, ‘indirect research’, ‘affective assessment’ or ‘using advanced queries’ were created as the respective behaviours were identified. Codes which did not generate many instances of similar behaviour, such as ‘image search’ or ‘Google fight’\(^{54}\), were not taken into consideration in the analysis. The open coding system facilitated the spontaneous observation of additional characteristics of research behaviours that became apparent with the larger group of participants. It later allowed the researcher to see whether these additional characteristics were displayed by more than one person and, thus, combined with other elements of the analysis, could be attributed to an element of a particular ‘style’. The quantifying of the qualitative data was also used to count other elements pertinent to the analysis such as the number of research units (RUs), the number of certain types of resources used in the task or the frequency of use of certain resources.

Last, but not least, codes were also used to establish various task-related and method-related aspects of the study such as the impact of the methods used in the study on the task performance.

Methodological and data triangulation

Triangulation is a means of using multiple data sources in an investigation to check the reliability and validity of the research evidence. Methodological triangulation in this study refers to the comparison of the same type of data from various parts of the study. For example, quantitative self-declared data from the Global Survey was compared with the same type of data from the Main Study. Data triangulation in this study refers to the comparison of two types of data (self-declared and observed) from the same study component, the Main Study.

Both types of triangulation were used in this study to obtain an answer to RQ2, the aim of which was to establish how much time on average translators spend on research activities during the translation process.

\(^{54}\) ‘Google fight’ is a technique that allows users to compare the number of search results returned by Google for two given queries.
(see section 5.2.1.) as well as to report on the use of types of resources and specific resources (see section 5.2.3. and 5.2.4.).

### 3.2.4. Instruments of data collection used in the present study

The following instruments were chosen to facilitate the collection of data for the different parts of the present study: an online survey tool, an online screen recording tool and a post-task questionnaire tool (see the unshaded areas of Figure 3.9). These instruments will be discussed below.

![Image](image.png)

**Figure 3.9 INSTRUMENTS OF DATA COLLECTION AS EMBEDDED IN THE OVERALL FLOW OF DATA**

**Online survey tool (Survey Monkey)**

Survey Monkey\(^{55}\) was chosen to collect data for the Global Survey and for the profile questionnaires accompanying the screen recordings in the Main Study. Survey Monkey facilitates not only the collection of data but also data analysis, including the possibility of cross-examination of the various results. It also allows the production of graphs, although the functionality of this component is limited. For this reason, the data gathered from the Global Survey was exported and analysed using Excel. It is also worth noting that an upgraded version of Survey Monkey was needed in order to conduct surveys of more than ten questions, to have flexibility during the design phase of the questionnaire and to access the more advanced analytical functions.

**Online screen recording software (Screencast-O-Matic)**

Screencast-O-Matic\(^{56}\) is a piece of software used in the present research to record all on-screen activities. It was used to collect audio and video data in the Pilot Study and in the Main Study. This software was chosen

---

\(^{55}\) https://www.surveymonkey.com/

\(^{56}\) http://www.screencast-o-matic.com/
over other well-known screen recording software applications such as Camtasia Studio because of its relatively low cost and the fact that it can be accessed online through a web browser. Participants can carry out the recording at their convenience with a minimum of set-up work required. They can also upload their files directly onto the Screencast-O-Matic’s server, where they can be accessed by the researcher. Screencast-O-Matic can be used as free software; however, for the purposes of this research, a Pro version had to be purchased in order to enable longer recordings (the free version only permits sessions of less than fifteen minutes to be recorded) as well as to give access to other useful features such as editing and script tools.

*Post-task questionnaires tool (e-mail)*

For the post-task questionnaire component of this study, a distance data collection method (a written questionnaire by e-mail) was chosen. It was used in the Pilot Study and in the Main Study. Both face-to-face and distance interviews were considered but, after evaluating the advantages and disadvantages of both methods, it was decided that for the purpose of this study a post-task e-mail questionnaire would be the best option for the reasons outlined in section 3.2.2.

3.2.5. **Instruments of data analysis used in the present study**

Statistical analysis tools and a mixed-methods analysis tool (shown in the unshaded areas of Figure 3.10) were chosen to support the qualitative and quantitative analyses in the present study.

![Figure 3.10 INSTRUMENTS OF DATA COLLECTION AS EMBEDDED IN THE OVERALL FLOW OF DATA](image)

*Statistical analysis tool (Survey Monkey, Excel)*

To aid the statistical analysis of the quantitative data, two tools, Survey Monkey and Excel, were used in the analysis of the data from the Global Survey and the Main Study. The analytical tools embedded in the

---

57 It is now possible to record screen directly from PowerPoint.
online survey tool Survey Monkey were used in order to gain a preliminary understanding of the data. However, due to their limitations, most calculations and graph production were performed in Excel, as already noted. In the Main Study, data was also imported into the mixed-methods data analysis software, MAXQDA, in order to interpret qualitative data in quantitative terms.

**Qualitative data analysis software with mixed-methods features (MAXQDA)**

MAXQDA is qualitative data analysis software with mixed-methods features which allows both qualitative and quantitative data to be imported and analysed in various ways. It can be used to organise, analyse, evaluate and interpret data and supports a variety of files ranging from text to multimedia files. In the present research, MAXQDA was used as the principal analytical tool during the Main Study in which data from the screen recordings, TAP, profile questionnaires and post-task email questionnaires was coded and quantified for the purpose of the analysis. Although MAXQDA was not used to its full potential in the present study, it proved very useful in coding, especially of the screen recordings.

### 3.2.6. Materials used in the present study

The materials used in the present study are the English source texts to be translated into the target languages of the participating translators. They were used in the Self-study, the Pilot Study and the Main Study.

In the Self-study, two source texts of different degrees of specialisation were used. Text 1, describing the eruption of the Sarychev Peak volcano, was taken from a popular science magazine *How it Works*. The more specialised Text 2, relating to the analysis of the components of the volcanic matter from the eruption of the Tungurahua volcano, was an excerpt from an academic paper which appeared in *Bulletin of Volcanology* (see Appendix 2). A comparison of the types of translation problems found in these two source texts and the consequent use of online resources to solve them was carried out, revealing predictable results. For example, it was noted that more time was spent on finding solutions to the problems found in the more specialised text and that more unpacking procedures, i.e. those related to source text comprehension, were followed in this text. This was to be expected given the nature of the research units, such as ‘sigmoidal particle density’ or ‘subplinian scoria fall deposit’, found in the specialised text. Therefore, it was decided to carry out the Main Study using only one source text in order to observe the differences between translators in relation to the same text. In other words, the focus was solely placed on translator research behaviour relative to one another, without the additional task variable of two different source texts. This decision was also motivated by what seemed feasible within the given time. Therefore, in the Pilot Study, the more specialised text was dropped and the text ‘The Sarychev Peak eruption’ used.

The reason for choosing one of the texts from the Self-study was motivated by the researcher’s familiarity with the translation problems present in the text and the search paths followed in the Self-study. The text was 245 words long and described some interesting features noticed during the eruption of a volcano. Potential translation problems which had been identified in the text during the Self-study included some volcano-related terminology such as ‘plume’, ‘caldera’ or ‘pyroclastic flow’, as well as proper nouns such as ‘Sarychev Peak’ or ‘Kuril Islands’. The text also contained a number of idiomatic and metaphorical expressions such as ‘to spark a heated debate’, ‘to mushroom’ or ‘capped with a head of steam’. It was
considered that the text contained a balanced amount of everyday language with potential translation problems that would prompt the participants to seek information and consult reference material. Readability scores were obtained for this text and, based on several formulas\textsuperscript{58}, the overall readability consensus was Grade Level 12, Reading Level: Fairly difficult to read, Reader’s Age: 17-18 years old.

However, after consultation with the participants in the Pilot Study, it appeared that the text did not pose a sufficient variety of problems, i.e. the translation problems present in the text did not prompt the participants to use a large variety of resources and, consequently, it was considered that a more demanding text should be used in the Main Study. Therefore, a text about Bitcoin and the emerging concept of digital currency (see Appendix 2) which spans the business/economic/financial domain was chosen. It was assumed that since the concept is relatively new, it would provide translators with ample information-seeking opportunities. The text contained interesting vocabulary such as ‘crypto-currency’, ‘fiat money’, ‘gold standard’, ‘hard currency’ or ‘peer-to-peer’ and some interesting ‘double-duty’ words such as ‘haircut’ or ‘mining’ which looked like promising candidates for prompting online research. The word ‘disruptive’ was also of interest due to its new recently-acquired meaning in relation to technology and innovation. Several named entities such as ‘WIR bank’, ‘Liberty Reserve’ or ‘Ithaca Hours’ were also considered to be potential candidates for instigating research activities. This is because, although they ‘sound’ like financial institutions, they are in fact alternative currency systems. A couple of phrases such as ‘wild swings’ and ‘the rise of’ were also of potential interest, as their translation in the context of digital currency would possibly require some adaptation. With the ecological validity of the study in mind it was important, however, for the text to be a naturally occurring one; not written or altered for the purposes of the study.

As with the previous text, several readability tests were performed and the overall readability consensus was established as Grade Level 14, Reading Level: Difficult to read, Reader’s Age: 21-22 years old, which suggested that the text was more difficult to read. However, the readability scores are used to assess the level of difficulty of texts for educational purposes and are not entirely suitable for assessing the difficulty of a text for translation purposes. They take into account the structure of the text but ignore the semantic or pragmatic dimensions. By comparison, excerpts from James Joyce’s \textit{Ulysses} or Friedrich Nietzsche’s philosophical works \textit{Thus Spake Zarathustra} and \textit{Beyond Good and Evil} were given much higher levels of readability indicating that they were much easier to read despite their well-known reputation to the contrary. Alternative ways of assessing the difficulty of a text for translation are also far from being adequate. Daems et al. (2015) used Lexile® levels which claim to provide “a more accurate representation of how challenging a text is”; however, Lexile® levels are still based on a quantitative approach and do not take into account the semantic properties of words and other aspects such as pragmatics, which play an important role in text comprehension and its subsequent rendition in another language. The model proposed by Reiss (Reiss, 1982, in Hale & Campbell, 2002:16), although a comprehensive one, is not empirically based and requires subjective evaluation. Although Hale & Campbell (2002) make an attempt at arriving at an empirically based way of assessing the difficulty of a source text\textsuperscript{59}, they conclude that “to this day there is not a single acknowledged procedure for grading texts used for translation teaching, or for

\begin{itemize}
\item Flesch Reading Ease, Gunning Fog, Flesch-Kincaid Grade, The Coleman-Liau Index, The SMOG Index, Automated Readability Index, Linsear Write Formula.
\item They do so by counting the number of different renditions of a particular lexical item posing a problem by a group of subjects, based on an assumption that the number of options is proportionate to the difficulty of rendering the particular item, i.e. the more options are available the more cognitive load associated with translating it and hence the bigger difficulty in translating it.
\end{itemize}
measuring the reliability of translation examinations” (Hale & Campbell, 2002:30). Although new approaches to grading text for translation are being developed, e.g. using Rhetorical Structure Theory ⁶⁰ used by Alves et al. (2010) to assess text organisation and coherence or text profiling by the PETRA group (Muñoz Martín, 2006) it could be said that still, no method has been widely adopted by the translation research community.

Although in the absence of other reliable methods readability scores were calculated in the present study, it is acknowledged that this method has severe limitations and is not particularly useful for assessing the difficulty level of a source text for translation. A new procedure based on empirical research would be most welcomed by translation teachers, examiners and researchers.

3.3. Research design

In order to answer the research questions of this study, a methodology not encountered before was crafted. This methodology, especially for dealing with the collected data, constitutes an added value of this study.

A flowchart representing the origin of data and the flow of this data from various methods of data collection and analysis to various instruments of data collection and analysis is presented below in Figure 3.11. Due to its complexity, the details of the Self-study and the Pilot Study are not included in the flowchart; however, they are described in sections 3.3.2. and 3.3.3. respectively.

The aims and the design of each of the study components will be described below, followed by a more critical engagement with the progression through the study components and how each of these components influenced the present study and its outcomes (section 3.3.5.).

⁶⁰ http://www.sfu.ca/rst/
3.3.1. The Global Survey

During the initial stage of this research, a general, large-scale preliminary data set was sought in order to make decisions regarding the focus of the Main Study and to help with the refining of the research questions. In the Encyclopaedia of Survey Methods, Lavrakas (2008: xxxv) observes that the survey method is used in research “to gather information to generate knowledge and to help make decisions”. He also points out the two primary characteristics of a survey: the fact that the sample is taken from the population (here: professional translators) and that each of the sampled members of the population is a source of data which is gathered by a systematic instrument, mostly a structured questionnaire. Gillham (2005: 24) also suggests using questionnaires “to carry out large-scale or a preliminary survey”.

Therefore, such a large-scale survey was deemed appropriate for this initial stage and it was conducted by means of an online questionnaire. The main purpose of this part of the research was to collect general information from a large number of the population of professional translators in order to ascertain their general habits and preferences with regards to the use of online search tools and resources and to determine some general trends such as:
- the average time spent on researching terminology
- the patterns of use of online versus paper resources
- the types of resources used by professional translators
- the degree to which the available resources are used by professional translators.

Figure 3.12 below shows the design of the Global Survey and the methods and instruments of data collection and analysis used.

![Figure 3.12 GLOBAL SURVEY DESIGN](image)

The online questionnaire used for the Global Survey was composed of sixteen questions (see Appendix 3). It was in circulation for four weeks, in the month of October 2012. The survey was disseminated under the name of Tools and Resources for Translation Professionals. The distribution channels included the ITI website, ITI Bulletin, ITI’s regional networks and forum groups such as ProZ and LinkedIn. Also, a number of responses came through personal contacts on LinkedIn, various mailings, blogs and Twitter. As noted, data for the Global Survey was collected using an online survey tool, Survey Monkey, which was then exported into Excel where most calculations were performed and graphs generated.

The questions were chosen to cover various aspects relevant to professional translators’ use of translation resources and covered a large selection of various technologies to look for frequencies of use. Mainly

---

61 https://www.linkedin.com/
closed questions were asked, and these were supplemented by comment boxes for additional information. Participants had the option to skip questions or their sub-parts as it was considered that a forced answer might skew the results, as noted in section 3.2.2.

The outcomes of the Global Survey were used to guide the further development of ideas and themes for a more in-depth investigation in the next phase of the study. One of the main findings of the Global Survey suggested that although translators reported spending a considerable amount of time on interacting with certain online resources (see section 5.2.1.), many existing resources are very much underused and translators are largely unaware of the wide range of resources available to them online. Another important insight was that research activities reportedly take the same amount of time regardless of whether translators use translation memory systems or terminology management systems. This led to the consideration of the role of online resources (as opposed to tools such as translation memory or terminology management) in the translation process and whether the way they are used is a contributing factor to the overall translation time. This question prompted further investigation into the interactions between professional translators and online resources. Due to the multi-dimensional aspects of these interactions, the analysis of individual cases was needed in order to answer this question, leading to the formulation of RQ3, which sets out to find out whether there are any distinctive patterns of behaviour (styles) with regard to the use of online translation resources.

3.3.2. The Self-study

The Self-study was carried out in December 2012 and was designed to self-examine the translation process of the researcher and the way online resources are used in this process. This was done in order to provide a starting point in the form of initial categories for the future analysis of translators’ interaction with online resources. These were then tested out in the Pilot Study and applied to the Main Study. Figure 3.13 below illustrates the design of the Self-study.

The Self-study was conducted by means of translating two texts of varying degrees of difficulty in MS Word while audio-recording comments about the process (TAP) by the researcher. As mentioned in section 3.2.6., the texts were taken from the domain of volcanology and were excerpts from a popular science text and an academic publication. During the subsequent analysis of the recordings, flowcharts were produced for each of the texts to capture the research activities undertaken during the translation process. In the flowcharts, translation problems were highlighted and the path leading to the solution recorded in writing. This provided some preliminary categories of analysis, according to which the translation process of the participants in the Pilot Study (and then in the Main Study) would be analysed. These categories included:
- total time devoted to research vs. translation activities
- number of units researched
- number of steps undertaken to find a solution to a problem
- number of resources consulted to find a solution to a problem
- types of knowledge sought: word vs. world knowledge.

Although no patterns could be detected in the Self-study for the obvious reason that data was based on one person’s performance, it was already apparent that a large proportion of the translation time was taken up looking for solutions to translation problems, navigating through various resources and checking or confirming a hypothesis in the target language, and also that a large number of resources were accessed during research activities with frequent repetitions. It was further observed that different sets of resources were consistently used for comprehension and reformulation procedures. These observations provided an initial framework for analysing the results of the next study component, the Pilot Study.

3.3.3. The Pilot Study

The Pilot Study was carried out in February/March 2013 and its aims were twofold. Firstly, it was conducted in order to deal with the procedural aspects and technicalities surrounding the Main Study, including testing of the screen recording software, assessing the suitability of the source text, testing the design of the profile questionnaire and the post-task questionnaire. Secondly, it was carried out in order to analyse the data gathered from the two participants according to the categories that emerged from the Self-study, to assess the suitability of these categories for use in the Main Study and to see whether any differences with regard to the research style of the two participants could be observed. Figure 3.14 below shows the design of the Pilot Study and the methods and instruments of data collection and analysis used.

The Pilot Study was carried out by two colleagues with similar personal profiles, i.e. both females in their late 20s with three to five years of experience and with similar backgrounds, having completed MA’s in translation at a UK university in recent years and one of them having completed a PhD in Translation
Studies. Katarzyna trained as an English-Polish translator and currently works in a translation company as a project manager. Saskia trained as an English-Italian translator and currently holds a Post-doctoral position. Both participants translate in a professional capacity, although translation is not their main occupation. It was not deemed necessary for the purpose of the Pilot Study to use experienced, professional translators for whom translation is their main occupation, as testing of the methodology and ironing out any technical issues were the main objectives of this study component.

This study component was carried out remotely and, as seen in Figure 3.14 above, consisted of profile questionnaires, screen recordings with audio commentary and post-task questionnaires. After the profile questionnaires were completed, detailed instructions and help sheets for the most challenging part of the task, the screen recordings, were sent out to the participants to help with accessing and using the software and to provide instructions about what was expected of them during the task (see Appendix 4). This was followed by post-task questionnaires which were delivered and returned via e-mail. The task to be recorded was based on an English source text from the domain of volcanology from a popular science magazine (see section 3.2.6. for details) and both participants translated into their L1. No brief was provided and this was critically noted by one of the participants and later included in the instructions to be given to the participants in the Main Study.

Despite the fact that only two participants took part in the Pilot Study, the analysis revealed significant differences in their research behaviour during the translation process. Firstly, they differed significantly in both the overall time of task completion and in the time devoted to the actual translation and research. Katarzyna took over twice as long to complete the translation, averaging just over 200 words per hour, with Saskia averaging around 450 words per hour. Based on the industry average of 2.5k words per day (300 word per hour), Katarzyna seems to come below this average and Saskia well above it. When it came to the ratios between the actual translating time and the researching time, Katarzyna’s observed time (35%) was within the 31-40% for unknown domains, which she self-declared in the questionnaire and which was in keeping with the average reported in the Global Survey. However, the amount of time spent on terminology research in unfamiliar domains (31-40%) self-declared by Saskia did not match the observed time, which was approximately 50%. The timings were calculated manually from the screen recordings and the time spent on terminology research was calculated whenever there was a switch from the translated text to an online resource.

The quantitative analysis encompassed the following categories:
- duration of translation phases (orientation/drafting/revision)
- number of resources used
- types of resources used
- proportion of time spent on translation activities vs. research activities
- types/frequency of actions (looking up, modification, confirmation)
- number of research episodes
- number of steps per research episode
- research techniques and frequency of their application (top-down vs. bottom-up).

---
62 As mentioned in section 2.6, data obtained from the Main Study participants was incomplete in this respect and therefore this analytical category was not pursued in the Main Study.
63 As above
The qualitative aspects of the analysis were:
- degree of familiarity with translation problem by each query (new, partially new, known)\textsuperscript{64}
- word vs. world knowledge (classification of translation problem according to whether the problem is a linguistic one or whether it relates to conceptual knowledge about the world)\textsuperscript{65}.

Generally, the following initial observations were made with regard to Saskia’s research style (reported here because the outcomes of the Pilot Study feed into the design of the Main Study described below):
- quick decisions
- intense use of internal resources
- relatively small number of research episodes (half Katarzyna’s number of episodes but episodes lasted much longer)
- the use of more top-down techniques, i.e. starting with a keyword and then navigating to a resource
- more emphasis on world knowledge - uses more knowledge-based resources such as encyclopaedias and Wiki.

Katarzyna’s research style can be characterised by the following initial observations:
- more deliberation with regards to decision-making
- more resort to external resources (compared to Saskia)
- a lot of actions taken to resolve translation problems, but the research episodes are short
- the use of both top-down and bottom-up techniques in equal measures
- more emphasis on word knowledge; uses more termino-lexical resources than Saskia
- cautious attitude - double-checking many known words (evidenced in audio-commentary).

Both Katarzyna and Saskia used roughly the same number of resources, but Katarzyna had twice as many research episodes. However, nearly a third of Katarzyna’s research episodes led to modifications, compared to one sixth in Saskia’s case. This would confirm the overall observation that Saskia is more confident in her choices and does not resort to the use of online resources as much as Katarzyna. This could also be confirmed by the fact that during the revision phase Saskia revisited two segments ‘marked for later’ and revised them without using any external resources.

Katarzyna, on the other hand, was generally more deliberate and cautious (nearly twice as many ‘checking/confirming’ actions as Saskia) and in her TAP she frequently used the word ‘check’, which indicates this cautious attitude. She was also much more concerned with applying the correct style and made numerous style corrections, whereas Saskia did not seem to consider this as much, as she did not make many stylistic adjustments and did not comment on style during the TAP.

The Pilot Study confirmed that the two translators differ significantly with regard to the amount of time they spend on research and how they conduct translation-oriented research. As will be seen in the analysis of research behaviour of the sixteen participants in the Main Study, most aspects of the research style found between the two Pilot Study participants were adopted in the Main Study; however some aspects, such as the confidence level or the speed of decision making, were not. This is due to the fact that additional, introspective insights would be needed to study these aspects in a larger sample and because of time limitations, this was not possible at this stage. Also, some aspects as noted in footnotes 60-63 could not be analysed due to incompleteness of obtained data.

\textsuperscript{64} As above
\textsuperscript{65} As above
It is also worth mentioning that the development of the various categories against which the data was analysed leading ultimately to the identification of a range of patterns of interaction with online resources was subject to the most significant evolution unfolding through the Self-study, Pilot Study and the Main Study. This evolution was a natural progression based on the increased amount of data obtained from a larger number of participants (1, 2 and 16 in the Self-study, Pilot Study and the Main Study respectively) and is underpinned by the basic tenets of grounded theory based on progressive identification and integration of categories of meaning from data.

3.3.4. The Main Study

The Main Study was conducted in 2013/2014 and the process of data collection took nine months, from July 2013 till March 2014. As mentioned in section 3.2., data collection proved to be one of the most challenging tasks in this study. Finding participants was problematic as no compensation for their time was offered and the study required a significant time investment in order to complete all parts of the Main Study (the profile questionnaire, the screen recording with audio commentary and the post-task questionnaire). It was originally envisaged that a sample of twenty would be representative enough and at the same time manageable within the time limitation of this study; however, due to the difficulties mentioned above, the sample was limited to sixteen translators. The participants were found through many channels, including newsletter announcements of professional bodies such as the ITI, various social media posts such as LinkedIn, Facebook and Yahoo groups, individual campaigning at conferences and personal channels of communication.

The biggest challenge in the Main Study was posed by the screen recordings, as setting up the software, uploading the files and dealing with any technicalities related to the delivery of the files was in the hands of the participants. Other difficulties with data collection included participants not delivering data as promised or delivering data that was unusable or not suitable for the purpose of the study. In the light of these challenges, it is understandable that translation process research focussed on professional translators can prove to be an overwhelming task that not many solo researchers are willing to undertake. It would, therefore, be of interest to TPR researchers to share their data, which could then be utilised for future research, as Göpferich (2010) suggests. Although a database of TPR data already exists (CRITT TPR-DB) (Carl, 2012a), it is restricted to keylogging and eye tracking data and, to the best of my knowledge, no database that would enable the sharing of various types of TRP data exists.

The design of the Main Study included four components: profile questionnaires, screen recordings accompanied by TAP and post-task email questionnaires (see Figure 3.15 below).

---

66 Although the categories in this study do not have a social dimension in the traditional understanding of the concept of ‘social interaction’, i.e. how humans relate to the world they live in, it could be said that they represent a new kind of social dimension, i.e. a digital one and relate to human-computer interaction (HCI), or even to human-information interaction (HII) (Zapata, 2015).
The methods of data collection and analysis were the same as in the Pilot Study, with the exception of coding and the mixed-methods data analysis tool MAXQDA which were employed in the Main Study in order to deal with the large amount of data generated that had to be analysed qualitatively and quantitatively.

Once the data was collected, the data processing started by manual and semi-automatic calculations, annotating and coding of the data. The screen recording for each participant was manually annotated in the form of a description of what a participant was doing or saying at the time as shown in the video image and three types of activities (translation, research and commenting) were identified. Each instance of each type of activity was logged and their identification allowed an Excel formula to calculate the distribution of translation, research and commentary activities for each participant. The following were the calculations performed for each participant:

- observed total recording time (hh:min:sec)
- observed translation activities (hh:min:sec)
- observed translation activities (% of the translation + research time)
- observed research activities (hh:min:sec)
- observed research activities (% of the translation + research time)
- observed comments activities (hh:min:sec)
- words translated per minute
- average translation episode (hh:min:sec)
- average research episode (hh:min:sec) (research time/number of research episodes)
- average research step (hh:min:sec)
- number of research steps per research episode
- number of research episodes
- number of research units (RUs)
- average time spent on an RU (hh:min:sec)
- number of specific resources used in the task
- number of types of resources used in the task (see section 4.3.4. for types of classifications used).
During the annotation process, all instances of resource use were noted and flowcharts including research sessions, research episodes and research steps taken by each of the participants were created (see Appendix 5).

To summarise, although the methods and instruments of data collection and analysis for this study were designed *a priori*, the actual methodology for analysing the individual differences, leading to the formulation of two taxonomies was developed *alongside* the process of analysing and organising the data as it was not anticipated at the outset what kind of patterns would be found. The progression from the Global Survey to the Main Study was an organic one and the methodological considerations and choices for the next step reflected the findings from the previous components. Figure 3.16 below summarises the present research from the research design to the hypothesis.

![Figure 3.16 THE STRUCTURE OF THE PRESENT STUDY - FROM RESEARCH DESIGN TO HYPOTHESIS](image)

### 3.4. Methodological challenges

During the process of study design, data collection, processing and analysing, a number of challenges were encountered. In this section, some of these challenges will be discussed.

**Establishing the duration of translation activities, research activities and commentaries**

As mentioned in section 3.2.2., difficulties related to the separation of translation activities, research activities and audio commentaries were encountered in the translation task of the Main Study. Due to the complex and often fluid nature of the intersections between translation and research activities, it would have been impossible to automatically separate translation activities from research activities. Therefore, this process was performed manually. This was also difficult at times, especially when longer pauses occurred due to the slow internet connection, or when it came to separating the audio commentary from the translation or research activities as this very often overlapped with either a translation or research episode. These challenges were met by calculating out the pauses due to technical reasons and other interruptions and the commentaries that were not overlapping the translation or research activities. When it came to deciding where the boundaries should be drawn in the cases of overlaps, best judgement was
exercised to delimit these boundaries. For example when there was a switch to a text and the participant was still talking but not actually translating, this time was calculated as commentary. Conversely, when there was a switch to a resource and the translator was talking whilst actually performing a search, this was calculated as a research activity. Additionally, research activities such as looking up synonyms in Word by right-clicking on a word also occurred within the text production window which is normally associated with the translation activities. This also had to be manually separated and included in the research timing. With the technology moving towards a more ‘blended’ approach (i.e. more resources being available from within the text production window) separating the translation and research time might become increasingly difficult. This has also been indirectly referred to by Muñoz Martin (2016:155) who points out that the increased use of technology might contribute to the blurring of the distinction between internal and external support discussed in section 2.1.

**Outliers**

Outliers in this study are those translators in the Main Study whose online information behaviour, or certain aspects of it, was observed to be far more extreme than that of the whole sample. Translators who displayed outlier behaviour were generally retained in the sample as they are “observations that represent a unique but valid aspect of the sample population” (Newton & Rudestam, 1999:172) and “contribute to a complete understanding of the phenomenon of the study” (Newton & Rudestam, 1999:169). However, the presence of outliers in a small sample such as the one in the Main Study can lead to distortions of statistical estimates and, consequently, to the inaccurate portrayal or distortion of the sample. Therefore, in calculations related to group behaviour and individual differences, outliers were treated on a case by case basis and were either included or excluded from the calculations, based on how far they were removed from the mean value for the whole sample.

These decisions were made by calculating z-scores. If a certain value is removed by 2.5 z-scores from the mean, it is considered to be an outlier (Newton & Rudestam, 1999:169) and that particular value is trimmed from the data thus narrowing the range. However, in this study, for the purpose of assigning the participant to a certain category, he or she was assigned to one of the ‘end of the scale’ categories (e.g., high or low), depending on which extreme end of behaviour he or she belonged to. For example, T1 was considered to be an outlier with her exceptionally high research time (see section 7.1.1.5.); and although her time was trimmed from the range, she was classified as having ‘High’ research time.

**The treatment of search engines (a tool or a resource?)**

Search engines were observed to be used in the Main Study in two ways: starting from a known source (these types of queries are known in information behaviour literature as informational queries (Broder, 2002) and starting from a query in a search engine which can be compared to navigational queries (ibid.). For example, if looking up the term ‘fiat money’ starting from a known resource, a translator would go straight to, for example, Linguee to see how it was translated before and possibly to check a definition. However, when starting from a query in a search engine, the translator would go to Google to type ‘fiat money’ and then look through the results to see if any of the links look promising. In this study, the former is referred to as a bottom-up strategy and the latter as a top-down strategy. Because in the bottom-up strategy a search engine could be seen as primarily helping to locate a suitable resource, rather than

---

searching for information, a question was posed as to whether search engines should be treated as tools to access resources on the entire web or as a type of a resource in itself.

As will be seen later, the distinction between tools and resources was an important one for the purpose of this study (details of how the boundaries were established are discussed in section 4.1.). However, the deciding factor in this particular debate was whether the use of Google, the only search engine used by the participants in the study, was applied to the whole translation task (in which case it would have been treated as a tool), or whether it was used ad hoc, at any time during the translation task to assist with individual research needs (in which case it would have been treated as a resource). Based on the fact that the use of Google was exclusively ad hoc, a decision was made to treat search engines as a resource. This decision was further based on the fact that the Google snippets were often consulted without translators actually clicking on the individual links.

Establishing research units

Research units were difficult to establish as it is not always clear which particular part of the text is being researched. In this study, research units were identified explicitly by means of audio commentaries (translators often uttered the actual term or expression under investigation) and screen recording (translators often copied the research units and pasted them into a resource). In the absence of such an explicit indication, the nature of particular research units was established based on the queries that were typed into a resource during the research activity. As mentioned in 3.2.3., as most of the 310 research units (82.3%) researched by the translators were explicitly indicated, all were given equal status as if they were explicitly indicated.

Enríquez Raído (2011:338) in her study of information behaviour of translation students and two professionals used online search reports in which participants reported their information needs. However, not all information needs were reported, which points to the fact that this method alone is also insufficient to establish all instances of an information need. Therefore, in future research, this difficulty in eliciting information pertaining to research needs and/or research units should be taken into account in research design and a more reliable methodology needs to be considered in order to establish these units with greater precision.

The instructions and technical issues

Both the Self-study and the Pilot Study provided important insights for the final design of the instructions given to the participants in the Main Study, especially with regard to the audio commentary (TAP). They revealed what questions needed to be asked in order to complement the screen recordings (e.g. providing information as comprehension vs. reformulation procedures, whether the word or phrase that was causing a problem was unknown, partially known or known but required checking, etc.). Unfortunately, despite the fact that instructions were given to all translators participating in the task, the TAP component was the least successful one in the study and yielded incomplete data as translators did not always comment on their procedures. Therefore, certain important aspects of research activities such as linking the use of particular resources to certain types of actions - looking up, modifying, confirming - could not be captured.

---

68 In the present study a distinction is made between a research need and a research unit (see section 3.2.3. for definitions).
69 The initial instructions were followed by a document containing ‘useful tips’ and a FAQ sheet (see Appendix 4).
One way of obtaining the missing information would be to ask the participants to recall such details during a retrospective viewing of the recorded translation; however, this was not possible due to the fact that the study was conducted remotely and although technologically feasible via the Internet, it would have been too challenging and time-consuming for both parties to consider.

In terms of the assessment of the technical aspects of carrying out the task, it appeared that although some difficulties were encountered in the Pilot Study, overall the setup and carrying out of the recording did not present many problems for the participants. Where minor difficulties did occur, feedback, mainly to do with troubleshooting, was provided by the participants with regard to the instructions. This feedback was incorporated in the final version of the instructions for the Main Study. Most participants reported no problems in setting up the screen recording software; however, three translators had issues relating to insufficient bandwidth or lack of expertise in using unknown software. Although most translators managed to upload their screen recordings onto the server, four participants reported having trouble with the uploading which resulted in frustration on their part. Nevertheless, most translators reported that the task was easier than they thought it would be and that they valued the experience of using screen recording software.

The technical issues surrounding process-oriented studies call for the development of a software component that could be connected to a translator’s existing translation environment, enabling recording of various types of data (key logs, web logs, screen recordings, eye tracking etc.) during the translation process without having to resort to external applications. Current tools used in TPR such as Translog are designed to research translation activities and as such do not allow monitoring of the use of external resources. Inputlog\(^70\), on the other hand, can record online activities as well as all keystrokes and mouse activities, but it cannot accurately log the use of translation memory or the use of machine translation.

\(^{70}\) http://www.inputlog.net/
4. Online resources for translators

The present work’s focus on the use of online resources by professional translators necessitates framing it in the larger context of translation technology. This is because translation technology has broadened beyond machine translation (MT) and CAT Tools (Computer Assisted/Aided Tools), also known as TEnTs (Translation Environment Tools) to include online resources, which were previously not always explicitly considered as part of translation technology.

In this chapter, firstly, the importance of asserting the position of online resources as part of translation technology will be discussed, followed by the consideration of the difficulties surrounding the definitions of translation technology and the terminological inconsistencies within its components. Also, a distinction between tools and resources will be drawn in order to justify the selected coverage of this particular section of translation technology in the present study.

Furthermore, because translation technology has diversified to embrace technologies from various neighbouring disciplines, it is important to frame it within the wider technological context to see what other technologies feed into it. This is in order to recognise the diverse nature of translation technologies and, more importantly for this study, the diversity of technologies that online resources draw on. Therefore, the second part of this Chapter considers the wider context which translation technology inhabits, and why it is important for professional translators to be aware of it. This context is expressed as two axes – a horizontal one where developments in other disciplines feeding into translation technology are considered, and a vertical one where the underlying architectures of technologies from each of these disciplines are discussed.

Finally, focusing solely on translation resources, various classifications will be discussed in order to systematise this vast and varied area of translation technology as well as for the methodological considerations relevant for this study.

4.1. Online resources as part of translation technology

In the past, dictionaries, encyclopaedias and other resources only existed in paper form, and, therefore, they were not considered as part of translation technology. It is only relatively recently that reference material consulted by translators has been subjected to the process of technologisation. The first CD-ROM version of the Oxford English Dictionary was only released in 1992 (OED:online) and its first online version was made available in 2000. The IATE\(^{71}\) term bank was launched in 2004 and made public only in 2007 (IATE:online). At first, this technologisation was confined to the simple digitisation of the content of the resources for online publication, but gradually the advances in computer science, language technologies and information and communication technologies were embraced by the developers of online resources, resulting in more technologically advanced solutions and thus bringing these resources closer to the realm of translation technology.

\(^{71}\) http://iate.europa.eu/
Although it is vital to acknowledge that “the relationship between translation and computers began with the development of software for machine translation” (Alcina, 2008:3), it is also important to recognise that the real boom in translation technologies started with the development of electronic dictionaries and terminology databases, the arrival of the Internet with its numerous possibilities for research, documentation and communication, and computer assisted translation tools (ibid.). The technological developments underpinning the architectures of online resources coincided with (and partly depended upon) the explosion of linguistic data on the Internet and the shift from the static, one-way web 1.0 to the dynamic, two-way traffic of web 2.0. Thus, with the support of the new technologies and the adoption of new trends enabled by technology, the digitised data was made accessible to translators on the Internet through a multitude of new ways, with faster and better targeted search facilities, more sophisticated and accurate information retrieval and new, collaborative ways of producing and sharing online resources.

Therefore, it is important to think of online dictionaries, corpora, term banks or even translation discussion fora, as belonging to the realm of translation technology, not only in terms of their function, i.e. extending the human capacity to remember, retrieve or find information, but also purely because of their architecture based on complex technologies. This way of thinking about online resources can already be observed in the recent literature. For example, Bowker & Corpas Pastor have included search engines, multilingual electronic dictionaries, corpora and concordancers “under the umbrella of the term of translation technologies” (Bowker & Corpas Pastor, 2015:12, original emphasis). Although this sounds very promising, more work is needed within the research and industry communities to assert the position of online resources within the remit of translation technology.

Unfortunately, in the current state of affairs, just as research is not always seen as an essential part of the translation process (as argued in section 1.1.), resources used by translators to carry out such research are not always explicitly associated with translation technology. The reasons for that could be many, but one of the main obstacles to a coherent vision of translation technology is the lack of up-to-date definitions of translation technology representing the field in its entirety as well as the inconsistent terminology used for some of its components.

Surprisingly, up-to-date definitions of translation technology are hard to come by. Even then, because they are often outdated, they are not always helpful. For example, Chan’s definition of translation technology as a “branch of translation studies that specializes in the issues and skills related to the computerization of translation” (2004:258) is somewhat vague and misleading in that it is confined to being a branch of translation studies. Bowker refers to translation technology as Computer Aided Technology which covers “any type of computerised tool that translators use to do their job” including Internet tools such as word processor, e-mail or World Wide Web and other tools such as corpus-analysis tools or terminology management systems (Bowker, 2002:6). Although Bowker does mention ‘Internet tools’, given that her book was published in 2002 when not many resources were available online72, online resources were not explicitly listed as part of translation technology. A similar picture is painted by Enríquez Raído & Austermühl (2002:2) where a typology of translation and localisation technology is presented, but with no definition of what this technology is, apart from making a distinction between ‘translator tools’ and ‘localizer tools’ seen as “separate, yet closely interconnected sub-processes”. Surprisingly, even the 2015

---

72 Although translation-related newsgroups, online forums and mailing lists which preceded social media were available from mid-1990s (Gaspari, 2015:580-581).
Routledge Encyclopedia of Translation Technology does not provide a clear definition. Instead, according to the editor, translation technology is defined by the scope of the Encyclopaedia which covers “all the concepts in the field and all the changes that translation technology has brought” (Chan, 2015:xxvii). Similarly, TAUS (Choudhury & McConnell, 2013) do not define translation technology as such, but propose that translation technology consists of Tools for the Professional Translation Industry and Machine Translation, the former being solutions that service the professional translation industry and the latter being key to the integration of translation into enterprise systems.

Therefore, given that “translation technology has become a norm in translation practice, an important part of translation studies, a new paradigm of translation pedagogy, and a major trend in the industry” (Chan, 2015), it is somewhat surprising that no adequate definition of translation technology reflecting the current developments has been offered. This lack of definitions is often reflected in translation technology being either confined to the realm of CAT Tools (Computer Assisted/Aided Translation Tools) or mistaken for MT (Machine Translation). This can be seen in Figure 4.1 below which shows a Google results page for the query ‘what is “translation technology”’. The prime example of the narrow, populist view of translation technology is the first result, a Guardian article in which all the ‘10 things you need to know about translation technology’ refer to various uses of MT (The Guardian, 2014:online).

![Google Search Results](https://www.taus.net/)

Figure 4.1 A SNAPSHOT OF GOOGLE SEARCH SHOWING DIFFERENT REPRESENTATIONS OF TRANSLATION TECHNOLOGY

---

73 https://www.taus.net/

74 An almost identical list was returned for query ‘define: “translation technology”’. 

---
Furthermore, the lack of consistent terminology describing the various components of translation technology also contributes to the terminological confusion. For example, CAT can be referred to as Computer Aided Translation or Computer Assisted Translation, with no differentiation between the two. Although the origin of CAT as an aftermath of the first disappointing implementations of machine translation, i.e. the move from FAHT (Fully Automated Human Translation) to MAHT (Machine-Aided Human Translation) is quite clear, the term ‘CAT Tools’, a much later designation of MAHT, has since become very broad and imprecise, covering many more tools than the original MAHT concept of translation memory with which it is often solely associated. The many blurred boundaries and unclear distinctions between some terms that describe translation technology tools, e.g. between CAT (computer-aided/assisted translation) and MAT (machine aided/assisted translation) are also emphasized by Quah (2006:6-8). To add to this confusion, the terms ‘translation tools’ or ‘translation memory tools’ are also often used as a substitute for CAT Tools, as are terms such as ‘translation software’, ‘machine-assisted translation’ or ‘machine-aided translation’. Zetzsche (2014a: 190-191) makes an attempt to bring some order into the disarray within the CAT Tools terminology by introducing a term TEnT (Translation Environment Tool) as a substitute for the term ‘CAT Tool’ in order to broaden the perception of CAT Tools beyond the function of translation memory. Bowker & Corpas Pastor (2015:3) list the following components which can be integrated into a TEnT: active terminology recognition, bitext aligner, concordancer, document analysis module, machine translation system, project management module, quality control module, term extractor, terminology management system and translation memory system.

Terminological inconsistencies can also be found with regard to online resources. The many ways these are referred to include: ‘translation aids’, ‘translation tools’ (Grego, 2010:110-116), ‘translator/knowledge tools’ (Enríquez Raído & Austermühl, 2002:2), ‘web-based resources’, ‘Internet resources’, ‘reference material’, ‘external resources’ (Zetzsche, 2015a), information resources (Zapata, 2015) external resources (Daems et al., 2015) or even ‘online ecosystem’ (Gaspari, 2015). Machine translation, too, suffers from being often mistakenly associated only with Google Translate, with customised solutions being frequently completely overlooked.

In view of the above-mentioned difficulties and inconsistencies, it seems important to provide a working definition of translation technology and establish what constitutes translation technology in the context of the current technological developments. Based on the proposed strengthening of the position of online resources within translation technology, it is argued that there are two main components of translation technology, namely tools and resources. Translation tools encompass machine translation and other tools commonly referred to as CAT Tools or TEnTs, such as translation memory, terminology management, alignment, extraction etc. and translation resources comprise a variety of technology and/or data applications such as dictionaries, term banks or corpora. Figure 4.2 below illustrates this distinction.
Alcina also makes a similar distinction. She considers *tools* to be “computer programs that enable translators to carry out a series of functions or tasks with a set of data that they have prepared and, at the same time, allows a particular kind of results to be obtained” and *resources* as “all sets of data that are organised in a particular manner and which can be looked up or used in the course of some phase of processing” (Alcina, 2008:94). In their typology of translation and localisation technology, Enríquez Raído & Austermühl (2002:2) also seem to support a similar distinction, however, they refer to both tools and resources, as ‘tools’. Thus, their ‘Localizer/Productivity tools’ which “aim primarily at streamlining the business process of translation” can be associated with what in this study is referred to as *tools*, and their ‘Translator/Knowledge Tools’ which “aim at providing the translators with the linguistic, encyclopedic, and cultural information” can be related to what is here referred to as *resources*.

Based on this distinction, it is proposed that translation technology is defined as *tools and resources used to aid, optimise or automate the translation process*. This definition, and more importantly, the stress on the treatment of resources as part of translation technology (although made for the purpose of the present work) might be useful in broadening the general understanding of what translation technology comprises.

However, like all attempts at classifications, this one is not without its own problems. As various technologies merge and complement one another, the boundaries between tools and resources are increasingly blurred (see Figure 4.3). These boundaries are fuzzy not only between the tools and resources but also within the two groups.
For example, MT is increasingly being integrated into CAT Tools/TEnTs as are some online resources, e.g. IATE is being incorporated into SDL Trados Studio\textsuperscript{75} (Maslias, 2015) and BabelNet\textsuperscript{76} was incorporated into XTM\textsuperscript{77} (Zetzsche, 2015b). Similarly, corpora and dictionary content are very much a core element of MT systems and concordancers rely on parallel texts. Some resources are built specifically using data from other resources, for example, BabelNet incorporates many different resources as shown in Figure 4.4 below.

![Figure 4.4 BABELNET – A MULTILINGUAL ENCYCLOPAEDIC DICTIONARY AND A SEMANTIC NETWORK INCORPORATING 6 DIFFERENT RESOURCES](image)

Recently, a new translation environment ‘Lilt’\textsuperscript{78} has shown to successfully combine translation memory, machine translation and a lexicon in an all-in-one tool in which all three components interact, learn from each other and the translator’s input, providing translation suggestions based on a segment level.

The difference between tools and resources may not be clear even to the developers themselves; for example, as we read in the newly redesigned version of Linguee, launched in early 2015, “Linguee is a unique translation tool combining an editorial dictionary and a search engine with which you can search billions of bilingual texts for words and expressions” (Linguee, 2015:online, emphasis mine). However, the ‘tool’ in this instance is the back-end infrastructure that enables Linguee to provide what is essentially data. Therefore, Linguee be considered a tool that acts as a resource. Similarly, announcing a new product from Reverso\textsuperscript{79}, ‘Reverso Context’, Hoffenberg (2014:8) refers to the many components of a future-style modern dictionary such as an example-based dictionary, a concordancer, a search engine, a bilingual aligner and an analysing tool, implying that in future this type of dictionary will be “a tool that helps us find the most appropriate choice in a certain context, offering users easy access to meanings and translations of words and phrases, along with relevant examples of usage, etc.” (Hoffenberg, 2014:8, emphasis mine). As with

\textsuperscript{75} http://www.sdl.com/cxc/language/translation-productivity/trados-studio/

\textsuperscript{76} http://babelnet.org/

\textsuperscript{77} http://xtm-intl.com/

\textsuperscript{78} http://lilt.com

\textsuperscript{79} http://www.reverso.net/
Linguee, the components Hoffenberg refers to are the technologies that underpin this modern dictionary, which still essentially remains a resource.

As mentioned before, the distinction between tools and resources was vital for this study because of its being limited to the study of the resources. Due to the fuzzy boundaries between tools and resources described above, further assumptions had to be made in order to help with the decision making process about whether the use of certain technologies by the participants in the present study should be included in the analysis.

Therefore, as illustrated in Figure 4.5, from the point of view of their purpose in the translation process, tools are understood to be technologies serving many different particular purposes in the translation process, such as retrieving previously translated segments, organising and retrieving terminology, counting words, converting formats etc. Their primary modus operandi is that of a mechanism, most likely a piece of software. Resources, on the other hand, are understood to be those technologies that assist in research associated with translation activities. They are usually collections of data underpinned by often complex technologies that make it possible to retrieve the desired information by the user and, increasingly, technologies that make this data “flexible, reusable, connected, sustainable, and efficient […] [enough] to adjust itself to the changing needs” (Pearsall, 2013:3). Tools and resources often depend on one another. However, whereas tools need data to perform their function (e.g. a translation memory system needs previously translated segments in the form of translation memory or a bitext), resources can often exist as a standalone technology that is directly accessed during the information seeking process independently of tools (e.g. dictionaries, glossaries or term banks).

A further assumption made for the purpose of this study was that tools are those technologies that are applied to the whole translation task, whereas resources are those technologies that are used ad hoc, at any time during the translation task to assist with individual research needs. This distinction was helpful in certain cases, for example with dealing with the different ways of using machine translation by four

---

80 Translation task is understood to encompass all activities between the opening and saving of the translated document.
participants in the Main Study. Whilst two of the participants applied MT for the whole task from within their CAT Tool (for example SDL Trados or Wordfast Anywhere), three of them also used Google Translate for ad hoc research during the task. Therefore, the uses of MT as a tool, although mentioned in the study, were not taken into account in the quantitative analysis (see also section 5.2.6).

Certain translation aids such as concordancers, search engines or machine translation can be seen as a tools, but can also can perform a function of a resource. In this dissertation they are referred to as ‘tools acting as a resource’.

4.2. Online resources as part of digital technology (wider perspective)

In contrast to the previous section which highlighted the need to strengthen the position of online resources within the wider scope of translation technology, this section seeks to illustrate the even wider context of the digital technologies in which online resources exist. This wider context takes into account the horizontal axis (the various disciplines that contribute to the development of online resources) and the vertical axis, which highlights the hierarchical nature of developments in these disciplines (see Figure 4.6).

Figure 4.6 ONLINE RESOURCES IN A WIDER CONTEXT OF DIGITAL TECHNOLOGIES

Considering the horizontal context, translation technology, and consequently online resources, are made up of various technologies coming from fields such as Language Technologies (including Natural language Processing and Computational Linguistics), Information and Communication Technologies and Computer Science, which are largely interdisciplinary fields themselves. From the vertical, architectural perspective, we can say that translation technology, and consequently, online resources, are made up of various infrastructure technologies.
Therefore, the underlying architecture of online resources can be seen as a continuum between infrastructure technologies (Hartley, 2009:108) and their application for a particular purpose. For example, the multilingual search engine 2Lingual is underpinned by cross-language information retrieval (CLIR) applied to create a particular tool or resource, in this case, 2Lingual, which allows a Google web search to be performed in two languages simultaneously. So, the applied technology that we access is the visible product of the underlying infrastructure technologies, often from many disciplines. This applied technology constitutes tools such as translation memory, machine translation, corpus analysis tools, linguistic search engines and resources such as online dictionaries, electronic corpora, term banks and so on, which are then employed by translators in their everyday work. Therefore, when referring to ‘translation technology’ in this study, what is meant is applied technology, although it is important to bear in mind the many infrastructure technologies that underpin translation technology.

Whilst it is not paramount that the user knows everything about the underlying architecture of the tools and resources they use, it is crucial to have a degree of awareness of the main principles they employ, in order to determine how they can be used in the translators’ work. It is important to know how they work, where the data comes from and what benefits a particular application might bring to an individual, especially now that many technologies combine to produce more and more sophisticated applications. For example, a translator might need to know whether a dictionary they want to use is corpus-based or not. They might want to know whether a particular corpus they wish to use is fixed in time or whether it is a monitor corpus that is constantly updated. They might need to know whether the data for the corpus was harvested automatically or whether it was compiled by humans. Translators might also want to know that Linguee operates on the basis of specially designed machine-learning search algorithms which filter out the most valuable bilingual concordances from the web and are capable of improving autonomously (Linguee, 2015:online). Or that semantic technologies allow resources such as BabelNet to connect concepts and named entities into a large network of semantic relations or that cross-language retrieval technology is used in multi-lingual search engines such as 2Lingual.

The infrastructure technologies are many and they are continuously morphing into new resources that can be utilised by translators. In addition, one person’s visible technology can become another’s infrastructure technology and the blurred boundaries observed between tools and resources or within these two groups mentioned earlier (see Figure 4.3) can also be observed between the visible and the invisible technologies. For example, the Europarl corpus can be used as a standalone tool, but is also widely used to train statistical machine translation (Koehn, 2005).

The difficulties in recognising the hierarchical structure of translation technology can be seen in the recently published Routledge Encyclopedia of Translation Technology (Chan, 2015). In Part III entitled ‘Specific Topics in Translation Technology’, various infrastructure technologies (e.g. POS tagging, Information retrieval and Text mining, Segmentation) are listed alongside the visible technology applications (Terminology management, Translation memory, Translation management) as well as whole disciplines (Natural Language Processing) and related topics such as Controlled language or Editing in translation. Whilst it is convenient to aggregate such varied and hierarchically diverse issues under a vague

---

81 http://www.2lingual.com/
82 Finding out the architectures and technologies behind a tool or a resource often proves to be a very difficult task. Understandably, a technology is very often secretly guarded to prevent others from copying. More often than not, gaining even a superficial understanding of ‘how it works’ frequently proves almost impossible.
83 http://www.statmt.org/europarl/
When talking about tools and resources, it is important to emphasise that both translation tools and translation resources inevitably encompass generic computerised aids not specifically designed for translators but extensively used by them, such as project management tools, image editors, online communication tools or Internet resources such as parallel texts or search engines. These, although not part of translation technology, can be called tools and/or resources for translators. In this thesis, all resources used by translators, whether they belong strictly to translation technology or not, are being taken into account in the analysis of how translators use online resources.

Due to the unprecedented growth in the volume, scope and variety of contemporary translation technology applications, it is challenging for professional translators to integrate these tools and resources into their translation environment. This is partly due to the lack of integration of these components within one translation technology platform which is reflected in Zetzsche’s efforts to communicate the ideas of professional translators to technology developers (Zetzsche, 2015b) in an attempt to include all useful translation technologies under one platform. It is interesting that apart from the many TEnT tools that are already included in most CAT software offerings, the two other components mentioned by professional translators as being desirable and/or requiring a better/deeper/different integration are machine translation and, more importantly for the object of this study, online resources. As Zetzsche concludes, an application allowing access “to all required external resources, whether they are browser-based or found in digital dictionaries” are what translators are frequently asking for (Zetzsche, 2015a & 2015b).

However, in order for this integration to take place, research is needed to find out what types of resources are used by translators and how they are used. It is hoped that the findings of this thesis will contribute to a better understanding of the possible avenues for this integration. The next section will be concerned with some classifications that might be helpful when thinking of translation resources collectively, as a group of resources that fulfil the same basic function of satisfying translators’ research needs. Since this work’s focus is on resources for translators, classifications of tools will not be taken into account. The reader can refer to Melby (1982), Alcina (2008) or TAUS (2013).

4.3. Classifications of online resources for translators

Online resources and tools acting as resources abound on the Internet and are growing exponentially. For example, at the time of writing, Lexicool.com84, a directory of free online bilingual and multilingual dictionaries and glossaries, offers access to over 8,000 dictionaries and glossaries (4,000 were offered in 2005). In 2011, the corpus query tool Sketch Engine85 offered 191 pre-loaded corpora in 47 languages (Jakubiček, 2011). Four years later, in 2015, it incorporates over 400 corpora in 70 languages (Kovár, 2015). Many resources (or tools that can be used as a resource) have only recently reached the development stage that makes them of potential benefit to translators. Google Translate, WebCorp86, BabelNet or even

---

84 http://www.lexicool.com/
85 http://www.sketchengine.co.uk/
86 http://www.webcorp.org.uk/live/
Linguee are relatively new additions to the pool of resources (or tools acting as resources), i.e. have been developed or become of use to translators in the last 8 years.

As mentioned before, translation resources are becoming more and more technologised and because of the new technological solutions, they are morphing into new types of resources. For example, BabelNet is a “multilingual encyclopedic dictionary, with lexicographic and encyclopedic coverage of terms, and a semantic network which connects concepts and named entities in a very large network of semantic relations” (BabelNet:online). It is, therefore, an encyclopaedia, dictionary and ontology, all in one (see Figure 4.7 below).

Figure 4.7 BABELNET - A DICTIONARY, ENCYCLOPAEDIA AND ONTOLOGY
Despite these difficulties in classifying certain resources due to their increased multi-functionality, the main types that are used by professional translators can be distinguished, fifteen of which are listed in Figure 4.8 and discussed individually in section 4.3.1. below.

4.3.1. Main types of resources for translation-related research

In this section, the main types of resources used for translation-related research will be discussed with regard to their usefulness in translation-related research. This discussion will encompass a brief overview of each type of resource, including recent developments and directions in the development of these resources.

**Figure 4.8 TYPES OF RESOURCES FOR TRANSLATORS**

**Dictionaries**

Dictionaries in this study are understood broadly as a collection of all types of dictionaries, such as monolingual, bilingual, general, specialised, synonym, collocation, idiom, orthographic etc. The history of dictionaries dates back several millennia to the "early civilisations of Akkadian, Babylonian, and Greek Empires" (Pearsall, 2013:2). The first attempt at organising the English language dates back to 1582, when Richard Mulcaster’s *Elementarie* was published, containing a list of 8000 words with no definitions (British Library:online) and later, in 1604, Robert Cawdrey’s *A Table Alphabeticall*, which is considered “the traditional starting point” for the English language (Pearsall, 2013:2).

Since then, the long history of lexicography has seen remarkable developments, which dramatically accelerated with the advent of the Internet. Kernerman (2013:7) lists several trends associated with the changes that have taken place in the development of lexicography, amongst them:

- from print to digital
- from tangible to virtual
- from one dictionary for life to many simultaneously
- from one-size-fits-all to customised and personalised
These trends encapsulate the changes that on the one hand caused dictionaries to become ubiquitous, but on the other, to “lose their autonomous identity and disappear in language technology” (Kernerman, 2013:1). As dictionaries are becoming one of the core components of machine translation, word processors, search engines, learning aids and the like, their content is being incorporated in and applied to new forms and tools “that go beyond the dictionary as end-product per se” (Kernerman, 2013:1). This necessitates the rethinking of the traditional concept of a dictionary, from a rigid set of data to flexible and reusable content that can be adjusted to the changing needs of the industry and the new user behaviours (Pearsall, 2013:3).

For translators, these changes mean that they now have more choice than ever and there is no need for hefty investments in the subscription-based resources; free does not necessarily mean poor quality anymore. Also, these changes will enable translators to work more flexibly and resourcefully in future as the content of dictionaries becomes embedded in their translation environments in a customised and personalised way thus facilitating efficiency and eliminating the waste of time spent on shifting interfaces. Translators also benefit from the growing popularity of collaborative functions embedded in dictionaries, which contribute to the emergence of new sharing practices and are aimed at providing translators with access to the latest knowledge and information, previously difficult to obtain due to the lack of flexible and sufficiently fast updates to reflect changes in language.

On the other hand, the explosion of other types of resources contributes to the decentralisation of the position of dictionaries. Nowadays, dictionaries are not necessarily used as the main resource for professional translators; they are becoming just one of the many types of resources translators have in their collections of reference material. Pastor & Alcina argue that translators turn to other resources as “dictionary creation cannot keep pace with developments in specialized fields” and that they are “deficient in the lack of information they include, the speed of content update, and the limited ways of accessing contents” (Pastor & Alcina, 2009). However, for some translators dictionaries still remain a core resource and thousands of different types of dictionaries are available to translators on the Internet in a multitude of language combinations. Admittedly, some languages fall behind with the development of their language resources; however, rapid advances are seen for many lesser known languages which often coincide with the development of machine translation for these languages.

Glossaries

Glossaries are often lumped together with other resources such as dictionaries. However, unlike dictionaries which provide comprehensive information about a particular word, including spelling, pronunciation, definition, variations, part of speech, origin etc., glossaries normally contain just enough information to identify a particular term or give a target language equivalent in a given context. Whilst dictionaries can contain general vocabulary, glossaries are always restricted to listing vocabulary related to
a particular area. The most popular way of providing such concise information is by means of a definition and/or target language equivalent; however, picture glossaries, such as donut glossary in Figure 4.9 below are becoming more and more popular.

![Figure 4.9 DONUT GLOSSARY](image)

Glossaries can often be found in collections of glossaries, for example on ProZ, where users are encouraged to submit their own glossaries which can then be accessed by other users and where individual entries can be voted on. Another well-known glossary collection is Glossary Links\(^87\), an offering from the terminology coordination unit at the European Parliament which contains some 3000 collections, two of which contain 2500 glossaries alone. There are many other collections, organised either by language, such as the extensive index available on TranslatorsCafé\(^88\) or by subject, for example, Glossarissimo\(^89\) or other collections\(^90\).

A large selection of glossaries is also available on Wikipedia. It allows the exploration of the database through links, linking various glossary pages, eventually leading to Wikipedia entries. An example of such exploratory search is shown in Figure 4.10.

---

\(^{87}\) http://termcoord.eu/glossarylinks/

\(^{88}\) http://www.translatorscafe.com/

\(^{89}\) https://glossarissimo.wordpress.com/

\(^{90}\) http://lai.com/thc/glsubj.html or http://stommel.tamu.edu/~baum/ hyperref.html#glossaries
Glossary of mill machinery

This glossary of mill machinery covers the major pieces of machinery to be found in windmills, watermills and horse mills. It does not cover machinery found in modern factories.[1]
**Terminology databases**

The need for a systematic approach to terminology arose with the rapid developments in science during the 18th century and, consequently, the need to name scientific concepts, especially as science became internationalised. As Sager et al. (1999:1-2) note, the field of terminology was driven by scientists and later, also by engineers and technicians. It was only in the second part of the 20th century that linguists became involved in the field. Therefore, terminology databases have always been driven by developments in science and technology which “required not only the naming of new concepts but also agreement on the terms used” (Sager et al., 1999:2). With the development of computer science, terminology has become even more relevant. As little as sixteen years ago Sager aptly captured the relationship between computer science and terminology as moving in two directions— computer science assists and changes terminological activities and its methodology, and terminology helps research in computational linguistics. [Terminology] has become a key factor in knowledge modelling and is used to transfer knowledge and technology, since it is by means of terminology that structured thought is transmitted. Terminology is at the basis of knowledge, and as such it is necessary for its representation. Cognitive science and knowledge engineering need terminology in order to advance.

(Sager et al., 1999:161)

Sager’s observations capture the importance of terminological databases today; that terminology is at the core of language technologies in their entirety. Any entity, be it an organisation, a company or an institution and any field of knowledge and its sub-fields is bound by certain linguistic parameters within which it operates. These parameters are or should be captured in order to facilitate communication. The terms, expressions and set phrases representative of the particular entity captured in a terminology database ensure consistency within all communications and, in the wider context, represent the knowledge base of this entity.

Translators can access these terminological resources as users or they can create their own terminological resources. Wright & Budin (2001:878) distinguish between termbases and term banks, whereby a termbase is an individual database produced by a company, agency or an individual and a term bank is an institutional resource, usually funded, which is accessible by subscription. This distinction highlights the personal vs. institutional nature of a terminology resource, but nowadays fewer resources are subscription-based and thus various term banks, such as IATE are also accessible free of charge. Other publicly available term banks include resources such as TERMIUM91 or EuroTermBank92. There are also enterprise-specific termbases available to translators, for example, Microsoft Language Portal93 or language-specific, such as National Terminology database for Irish94.

Termbases are nowadays normally associated with terminology management systems and function as a terminology database integrated with a translation memory system. However, one of the largest term banks available online, IATE, also calls itself a ‘termbase’, which brings confusion to the terminology associated with terminographic resources. In this work, term bank has been used to refer to both

91 http://www.btb.termiumplus.gc.ca/
92 http://www.eurotermbank.com/
93 http://www.microsoft.com/Language/
94 http://www.tearma.ie/
Corpora

Compared to the hundreds of years of lexicographical tradition, corpus linguistics is a relatively new field, in which corpora were initially built for language research purposes. According to Lan (2015:465), “the landmark of modern corpora is generally attributed to the Brown Corpus of Standard American English” built in the 1960s and consisting of one million words. Since then, corpus linguistics has seen rapid developments which “can be directly or indirectly attributed to advances in technology” (Renouf, 2005:np). These developments, affecting the design, scope and purpose of corpora, are summarised by Renouf (ibid.) as follows:
- the growth from small to large corpus
- the shift from synchronic to diachronic corpora
- from the designed corpus to collecting large amounts of texts
- from single focus to multi-dimensional corpora.

Due to the lengthy process of adoption of corpora as a resource for translators, their use at the end of the last decade had “not yet become widely established among professional translators” (Aston, 2009: IX-X). Empirical research, too, shows that in translation practice corpora are a still very much underused resource (Gallego-Hernández, 2015; Gough et al., 2013:53). However, Pastor & Alcina (2009:13) argue that recently, translators have been increasingly turning to corpora in search of terminology due to the inadequate developments in the field of dictionary creation. With many corpora being easily accessible and free of charge, “translators are among those who have much to gain from using corpora” (Frankenberg-Garcia, 2015:1).

Corpora are understood to be large, principled and systematic collections of naturally occurring texts. There are many different types of corpora, for example:
- general/reference vs. specialised corpora
- historical corpora vs. corpora of present-day language
- monolingual vs. bilingual corpora
- spoken vs. written vs. mixed corpora.

The benefits of corpora for translators are many. Frankenberg-Garcia argues that corpora can help with many language questions “for which dictionaries, grammars and other language resources do not always have clear answers” (Frankenberg-Garcia, 2015). Monolingual corpora can be used to discover target language patterns such as grammar or collocations, or strengthen knowledge of these patterns. They can be used, too, to check or verify hypotheses in the target language, an activity that promotes a target-oriented research in translation. Specialised corpora can also be useful as a resource to help translators with specialised terminology and phraseology. Aligned parallel corpora (also called translation corpora) can be an effective resource for establishing equivalence due to the availability of the surrounding context. Also, as Frankenberg-Garcia (2014) observes, parallel corpora can raise translators’ awareness about changes that are required at discourse level, which promotes the adoption of more creative and less literal translation strategies.
Interestingly, many modern dictionaries are now based on corpora. In the past, dictionaries were ‘example based’, i.e. the meaning of a word was established by means of “examples of words in use, taken from books or other sources - as a basis for describing language” (Macmillan:online). Nowadays, dictionaries are increasingly being built using corpora, which allow dictionary developers to gain insight into how a word behaves based on a large body of text, rather than on one example. This development was made possible by the advancements in computer sciences, the explosion of data on the Internet and the development of methodologies enabling the harvesting and processing of this data for linguistic purposes. As Fantinuoli & Zanettin (2015:1) observe, “taking advantage of advancements in terms of computational power and increasing the availability of electronic texts, enormous progress has been made in the last 20 years or so as regards the development of applications for professional translators”. Certain corpus querying tools such as AntConc
95 are available as freeware, although the more sophisticated ones such as Sketch Engine are fee-based. However, the developers of Sketch Engine also offer a lightweight version of their software (SKELL), currently only in English (other languages to follow) which is available for free
96. Figure 4.12 below shows a comparison of the corpus query for the word ‘assignment’ in Sketch Engine and SKELL. In Sketch Engine, translators have numerous options of querying some 400 corpora in 70 languages as well as creating their own corpora from the web and their uploaded translation memories. In SKELL, however, users can only query the system for three basic queries -concordances, word sketches and related words - in one, specially designed corpus consisting of 1 billion words, including the BNC corpus
97 and texts from news publications, academic papers, Wikipedia articles, open-source (non-)fiction books, web pages, discussion forums, blogs etc. (Baisa & Suchomel, 2014:66). The lightweight nature of SKELL, albeit admittedly of limited use for professional translators, illustrates the trend towards making corpora more accessible to a wider audience.

Figure 4.11 SKETCH ENGINE

95 http://www.laurenceanthony.net/software.html/
96 https://skell.sketchengine.co.uk/run.cgi/skell/
97 http://www.natcorp.ox.ac.uk/
Therefore, with easy access to many large corpora, the increasingly easy and achievable ways of creating their own corpus-based resources and the development of corpus-query tools facilitating flexible ways of information retrieval, it is not difficult to see how translators might find it beneficial to use corpora for their translation research.

**Thesauri**

Although the origins of thesauri can be traced to Philo of Byblos in the 4th century, the first modern thesaurus was published in 1852 by Peter Mark Roget and, to date, is one of the most widely used reference works ever published. Roget’s was the first language reference book to be arranged by topic and the first to encompass the semantic network of the entire language (Hüllen, 2004). Named ‘Thesaurus of English Words and Phrases, Classified and Arranged so as to Facilitate the Expression of Ideas’, it listed words grouped according to similarity of meaning and lexical relations. This was in contrast to dictionaries which provided words listed in alphabetical order and their definitions. To quote Roget himself, the object of his work was precisely the opposite to that of a dictionary, i.e. “to find the word, or words, by which [an] idea may be most fitly and aptly expressed” (Roget, 1853:3). Roget’s hyperlinked thesaurus is now available online and provides various options for looking up word-sense relations; however, it does not appear to be particularly user friendly. Other online thesauri include Thesaurus.com and many online dictionaries also incorporate thesauri as part of their package.

Visual thesauri resembling mind maps are also available, for example the Thinkmap Visual Thesaurus, Graph Words, or Snappy Words which provide visualisations of how a word relates to other words around it (see Figure 4.13).

---

98 http://www.roget.org/
99 http://www.thesaurus.com/
100 http://www.visualthesaurus.com/
101 http://graphwords.com/
102 http://www.snappywords.com/
Concordancers

Concordancers are translation aids that can be seen both as a tool and a resource. As a tool, they are associated with corpus analysis tools and provide one of their main functions - concordancing. This function enables users to find all occurrences of a searched word in a corpus, together with a span of surrounding text, to the left and right, as shown in Figure 4.14. Therefore, they are essentially designed to interrogate corpora.
Concordancers as tools can also be found as part of translation memory systems, where they are used to manually retrieve matches from a translation memory. As this study is not concerned with tools, these aspects will not be considered here\textsuperscript{103}.

However, it is also possible to see concordancers as resources. For example Valli (2012), although referring to concordancers as ‘tools’, looks at them from a resource perspective. She gives examples of sites such as MyMemory\textsuperscript{104}, Glosbe\textsuperscript{105}, TAUS Data\textsuperscript{106} or Linguee, which are essentially online resources with functions for searching for previously translated texts. The collections of previously translated texts are often contributed by other users in the form of translation memories (e.g. MyMemory, TAUS) or harvested from the web (Glosbe, Linguee). Although the function of a concordancer is the same as that of corpus analysis tools, here, the data that is interrogated is a collection of translation memories rather than purpose-built corpora. Interestingly, both Glosbe and Linguee claim to be online dictionaries, which highlights the fuzziness of boundaries between resources as mentioned earlier and emphasises how new technologies, in this case sophisticated web-crawling, are changing the shape of resources.

In this study, concordancers such as TAUS Data or Linguee were classified as resources and included in the data analysis.

Knowledge-based resources

Knowledge-based resources in this work are understood to be all resources that contain solely encyclopaedic information such as encyclopaedias, Wikis, compendia, or information databases.

Much has been said about Wikipedia not being a trustworthy resource due to the collaborative nature of the production of its entries; however, the fact that it is invariably ranked in the top ten most visited websites in the world also makes it one of the top resources for professional translators. Wikipedia’s lesser known academic-style counterpart, Scholarpedia\textsuperscript{107}, is an example of a collaborative resource that can be fully trusted as it undergoes rigorous peer review, similar to that in academic peer-reviewed journals and can even be cited in academic writing.

Knowledge-based resources also encompass computational knowledge engines such as Wolfram Alpha\textsuperscript{108} which provide “a fundamentally new way to get knowledge and answers - not by searching the web, but by doing dynamic computations based on a vast collection of built-in data, algorithms, and methods” (WolframAlpha:online). Resources such as this differ fundamentally from search engines in that instead of providing a list of links to websites or documents that might contain an answer, they answer factual queries directly by computing the in-built ‘curated data’. Figure 4.15 below shows some examples of questions that can be submitted to WolframAlpha in the Words & languages category.

\textsuperscript{103} A selection of available corpus analysis tools can be found here: http://cw.routledge.com/textbooks/0415236991/links/corpus.html
\textsuperscript{104} https://mymemory.translated.net/
\textsuperscript{105} https://glosbe.com/
\textsuperscript{106} https://www.tausdata.org/
\textsuperscript{107} http://www.scholarpedia.org/article/Main_Page/
\textsuperscript{108} http://www.wolframalpha.com/
Figure 4.15 WOLFRAM ALPHA – EXAMPLES OF QUERIES IN THE WORDS & LANGUAGES CATEGORY

Figure 4.16 below shows the many other categories in which queries can be submitted for computation.

Figure 4.16 WOLFRAM ALPHA – CATEGORIES IN WHICH QUERIES CAN BE SUBMITTED FOR COMPUTATION BASED ON BUILT-IN DATA, ALGORITHMS AND METHODS
Ontologies

Ontologies are not immediately associated with resources for translators, although they can provide a valuable source of information, especially in the initial phases of research, to support domain conceptualisation. Bowker & Corpas Pastor (2015:13) suggest that ontological resources enable translators to obtain “a preliminary vision of the domain […] to identify core terms and multiword units, to assess the degree of concept correspondence between the SL and the TL\textsuperscript{109} terms and to establish other potential equivalents”.

A domain ontology is essentially a “complex data structure that introduces formal concepts and describes the relations existing between those concepts” (Declerck et al., 2006:1492). Although designed to be read by machines and mostly used in semantic web and NLP applications, ontologies can be a useful resource for translators although stand-alone ontologies that can be directly accessed by individuals are scarce or have for some reason disappeared from the Internet. For example Mindpedia, which provided conceptual maps of words or terms based on Wikipedia articles, is no longer available on the Internet, although it was operational as recently as 2013. Certain resources, however, incorporate ontology into their design. A good example is BabelNet, which was mentioned in section 4.1.

Although ontologies have not yet found their way into translator toolboxes, it is very much possible that future applications will allow valuable resources such as YAGO 3\textsuperscript{110} to be made available to translators either through a standalone application or embedded in their translation environment.

Search directories/knowledge portals

Search directories can be called ‘human search engines’ as they contain a collection of links to web pages that are selected and organised by humans (Fransson, 2010:29). The content of a directory has to be structured and curated. According to Fransson (2010:30) links included in search directories have to be: well written, created by authoritative originators, stable (not likely to move or change contents), of good quality and kept up-to-date. Therefore, compared to the collection of links automatically selected by a crawler-type search engine during a search process, search directories can be regarded as a high-quality resource.

As mentioned before, search directories are becoming obsolete (Fransson, 2010:30). This is because human-powered resources that need constant maintenance are demanding in terms of time, effort and funding. The closure of Yahoo! Directory in 2014 and the closure of another highly regarded search directory with a long, twenty-year history, Librarian’s Internet Index\textsuperscript{111}, could be seen as a herald of this type of resource being superseded by alternative solutions. These new solutions, rather than being a collection of links to websites, are based on providing specific information that the user asks for in the form of knowledge portals, also known as knowledge sharing services or online information services. This shift supports Liddy’s (2001:np) observation that “consumers would prefer a finding engine rather than a search engine”.

\textsuperscript{109} SL refers to the source language and TL to target language.
\textsuperscript{110} http://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/yago/
\textsuperscript{111} http://www.ipl.org/
Taken at face value, the definition of knowledge portals as “gateways to information and other resources available from various sources and sites on the Internet or intranets, extranets or other internally stored databases and systems” (Toolbox, 2008:online) could pass for a definition of a web directory. However, there is a big difference between the two: whilst web directories provide links to valuable content available on the web, knowledge portals claim to provide valuable content that they create and curate themselves. They do not just provide links to websites, but “support knowledge creation, production, acquisition, aggregation, filtering, organization, transmission, dissemination, usage and/or retention” (Toolbox, 2008:online). Examples of knowledge portals offering general content include About.com\textsuperscript{112} and Mahalo\textsuperscript{113}. In the specialist domains, good examples would be portals such as Eldis\textsuperscript{114}, a portal for global development knowledge or Scidev.Net\textsuperscript{115}, a portal offering content related to science and technology for global development.

To summarise, web directories and knowledge portals are useful for translators searching for a broad subject or idea and wanting to see only relevant pages recommended by experts (Fransson, 2010:30). Such resources could be useful for domain conceptualisation and acquiring general knowledge about a subject before embarking on a translation project in an unfamiliar domain.

### Discussion fora

Discussion fora are centred around online communities of translators and serve the purpose of enabling discussion of various work-related issues. They started to appear in the 1990s, first as e-mail enabled mailing lists and then as usenet-based groups, supported by “rather rudimentary internet-based communication facilities” (Gaspari, 2015:579). With the evolution of the web from web 1.0 to web 2.0, these groups evolved into discussion fora, which “had additional user-friendly features designed to make the exchange of posting smoother” (Gaspari, 2015:580). Although Gaspari argues that discussion fora are being replaced by social media sites such as Facebook\textsuperscript{116}, LinkedIn or Twitter\textsuperscript{117}, this is not necessarily true for research-related questions. Discussion fora such as those on ProZ, WordReference\textsuperscript{118} or TranslatorsCafé serve as valuable sources of information and as ‘consultation rooms’ for peers who often provide expert advice on terminology-related questions or indeed, any other subject. They are often very organised, with subject searchable, specific directories and peer-reviewed entries. Furthermore, there are many language-specific discussion groups or groups dedicated to certain aspects of translation, for example, GP Search for sharing web searching techniques\textsuperscript{119} or GlossPost for sharing new glossaries or online dictionaries\textsuperscript{120}. A list of various discussion groups and mailing lists is available on translationjournal.net\textsuperscript{121}.

\textsuperscript{112} http://www.about.com/
\textsuperscript{113} http://www.mahalo.com/
\textsuperscript{114} http://www.eldis.org/
\textsuperscript{115} http://www.scidev.net/global/
\textsuperscript{116} https://www.facebook.com/
\textsuperscript{117} https://twitter.com/
\textsuperscript{118} http://forum.wordreference.com/
\textsuperscript{119} https://groups.yahoo.com/neo/groups/gpsearch/info/
\textsuperscript{120} https://groups.yahoo.com/neo/groups/GlossPost/info/
\textsuperscript{121} http://translationjournal.net/journal/00disc.htm. or here: http://multilingualbooks.com/translation-discussion.html
**Search engines**

Search engines, like concordancers, could be perceived as tools or tools acting as a resource, depending on the search strategy adopted by the searcher. If we consider the entire web content as a resource in itself and the search engine as a means of accessing this content, then the search engine could be regarded as a tool. In this mode, the searcher knows what they are looking for and uses a search engine solely as a vehicle to get to the resource they require. However, if we consider the actual results pages as the content being consulted, which often happens when carrying out translation-oriented research, then a search engine could be regarded as a tool acting as a resource. In this mode, the search engine acts as a source of information in itself, where the searcher consults the search results and might or might not follow the links to resources. In this study, for reasons explained in section 3.4., search engines are treated as tool which act as a resource.

As Lewandowski (2012:50) aptly observes, “web search engines are an impressive feat of software engineering, incorporating technologies from a wide variety of fields of computer science”. They are the central starting point for explorations of web content, and yet “[e]ven though most users use search engines every day, they know very little about them” (Lewandowski, 2012:2). The history of search engines began in the early 1990s with Archie and Gopher, but the first true search engine which searched through all of the manually maintained lists of web resources was the W3Catalog, developed by Oscar Nierstrasz in 1993. Currently, there are around 250 different search engines on the Internet with one search engine, Google, dominating the market. Data from *Search engine market share* (online) reveals that, as of July 2015, Google provides 70.80% of search coverage, followed by Bing (9.83%), Yahoo Global (9.57%), Baidu (7.52%). The remaining players in the search engine market have less than 1% of market share.

Developments in web searching are fast-paced and search engines are subject to constant evolution. Therefore, it is a challenging task to capture the latest ‘state of the art’ search engine technology at any given time. As of 2015, in terms of the indexing techniques used in the process of information retrieval, there are three types of search engines (Ritika & Gaurav, 2013:1546):

- crawler-based search engines which create their listings automatically by ‘crawling’ the web and ranking pages algorithmically
- human-powered search engines in which listings are created and categorised by humans
- hybrid search engines which combine these two approaches.

However, the search directory model seems to be fading out and not many truly human-powered directories currently exist. The oldest running, and a pioneer in the field, *Yahoo! Directory* was closed in 2014, twenty years after it was launched (Sullivan, 2014). Instead, other forms of human-powered search engines are emerging in the form of Q&A engines or social search engines (Lewandowski, 2012:19-46). Due to the fast-paced developments in the field of web-searching, it is often difficult to tell which search engines operate on what principles at any given time but the most notable web-crawlers are Google and

---

122 A list can be found here: <http://www.thesearchenginelist.com/>  
123 According to Lewandowski in most countries Google has a market share of 90% (Lewandowski, 2012:3)
Baidu$^{124}$ and examples of directory search type search engines are Open Directory$^{125}$ and WebDirectory$^{126}$. Examples of hybrid search engines are not apparent from the literature covering this subject.

Search engines can also be distinguished from the point of view of the different Information Retrieval techniques (IR) used, for example:
- keyword search engines, which do not take context or meaning into account (e.g. Google)
- natural language search engines, which use Natural Language Processing techniques during the search process, usually in the form of Q&A (e.g. Ask Jeeves$^{127}$, Bing)
- semantic search engines such as Swoogle$^{128}$, Watson$^{129}$ or Sensebot$^{130}$ which crawl the web for rdf$^{131}$ data. These search engines, although in their nascent state, are designed to provide results based on the sense (linguistic data contained in rdf descriptions) rather than statistics.

Additionally, the new, collaborative trends have resulted in new forms of search engines:
- social search engines, e.g. Sproose$^{132}$ or ZEEF$^{133}$ which use the social networks of a person to retrieve relevant data based on these networks
- crowdsourcing-based models that employ either paid or community-based methods to answer questions (e.g. Ask Jeeves or ChaCha$^{134}$).

Figure 4.17, Figure 4.18 and Figure 4.19 below show results for the same query ‘what is fiat money’ in three different search engines: Google, Open Directory and ChaCha, showing the differences in displayed results based on the use of different models of information retrieval.

---

$^{124}$ http://www.baidu.com/
$^{125}$ https://www.dmoz.org/
$^{126}$ http://www.opendirectorys.com/
$^{127}$ http://uk.ask.com/
$^{128}$ http://swoogle.umbc.edu/
$^{129}$ http://watson.kmi.open.ac.uk/WatsonWUI/
$^{130}$ http://www.sensebot.net/
$^{131}$ Rdf (Resource description framework), is a framework for describing data on the Web
$^{132}$ http://sproose.com/
$^{133}$ https://zeef.com/
$^{134}$ http://www.chacha.com/
Figure 4.17 AN EXAMPLE OF A KEYWORD SEARCH ENGINE

- Google
- What is "flat money"?
- About 469,000 results (0.33 seconds)

**Fiat money** - Wikipedia, the free encyclopedia
- Fiat money is currency which derives its value from government regulation or law. The term derives from the Latin facer ("to be done," "it shall be"). It differs from commodity money in that it does not have intrinsic value.
- Commonly, it is also referred to as "paper money.
- Commodity money - Iraq Dinar - Money creation - Hard currency

**Fiat Money Definition** - Investopedia
- Currency that a government has declared to be legal tender, but is not backed by a physical commodity. The value of fiat money is derived from the relationship between supply and demand rather than the value of the material that the money is made of.

**How the Fiat System Works - For Dummies**
- www.dummies.com/how-to/dummies/how-to/fiatsystemworks.htm
- Most of the world’s money is called fiat money, meaning it is accepted as money because a government says that it’s legal tender, and the public has enough faith in the system.

**Flato Currency: A History of Failure - Daily Reckoning**
- dailyreckoning.com/flat-currency/

Figure 4.18 AN EXAMPLE OF A DIRECTORY SEARCH ENGINE

- DA MOZ
- Search: what is flat money
- Links to explore categories associated with the search query
- Links to websites containing the search query

**DA MOZ Categories**
- 2. Society: Income, Economic Issues, Money Sales, Tax

**DA MOZ Sites**
- 1. Translation for the Advancement of Monetary Education
- 2. Fiat - Money exchange for Bitcoin and Litecoin, but also offers services for several other cryptocurrencies and fiat currencies.
- 3. Fiat Flat: Bank
- 4. Fiat Flat: Bank: Information about the fiat flat, and how to use it.
- 5. Fiat Flat: Bank: Additional information about the fiat flat.
- 6. Fiat Flat: Bank: Additional information about the fiat flat.
- 7. Fiat Flat: Website: The official fiat flat website, with a dashboard, history, contact information, dealer listings, event calendar, and news.
There are many other ways of categorising search engines, for example, general vs. meta search engines. General search engines are single, crawler-based engines such as Bing and meta search engines (also called aggregators) are those that base their results on the results of other search engines (e.g. Dogpile\textsuperscript{135}, Metacrawler\textsuperscript{136} or Mamma\textsuperscript{137}). Another popular way of searching is by data type: audio (e.g. Last.fm\textsuperscript{138}), video (e.g. YouTube\textsuperscript{139}), images (e.g. Picsearch\textsuperscript{140}) or datasets (e.g. Quandl\textsuperscript{141}). Of special interest for translators, however, is the proliferation of specialised search engines providing a wealth of information in specialised domains, for example, Lexis\textsuperscript{142} for legal research or PogoFrog\textsuperscript{143} for medical information.

\textit{Web pages (parallel texts)}

Web pages are hypertext documents that are connected to the World Wide Web. A collection of web pages constitutes a website. Although from a technical point of view any online resources can be considered to be websites, here we take a functional perspective and look at the resources not from the point of view of their architecture but from the point of view of their usefulness for translators carrying out research.

\begin{itemize}
  \item \textsuperscript{135} http://www.dogpile.com/
  \item \textsuperscript{136} http://www.zoo.com/
  \item \textsuperscript{137} http://www.mamma.com/
  \item \textsuperscript{138} http://www.last.fm/
  \item \textsuperscript{139} https://www.youtube.com/
  \item \textsuperscript{140} http://www.picsearch.com/
  \item \textsuperscript{141} https://www.quandl.com/
  \item \textsuperscript{142} https://www.lexisnexis.com/
  \item \textsuperscript{143} http://www.pogofrog.com/
\end{itemize}
According to a daily estimator of the size of the web\textsuperscript{144}, as of 8 July 2015 there are at least 4.65 billion web pages on the Internet. As mentioned elsewhere, such is the nature of a translator’s job that in order to find terminology for their translations they sometimes consult the most unlikely of resources. Web pages are an excellent source of information that cannot be found in any of the resources specially crafted for linguistic purposes. They are used by professional translators as parallel texts from which terminology can be harvested. Additionally, more and more websites now include glossaries containing vocabulary relevant to the subject of the website. For example, the bitcoin.org website features a glossary of Bitcoin-related terms, three of which actually appeared in the source text used in the Main Study (see Figure 4.20 below).

![The 'Bitcoin' website containing a glossary of relevant terms](image)

**Figure 4.20** THE ‘BITCOIN’ WEBSITE CONTAINING A GLOSSARY OF RELEVANT TERMS

But websites are used not only as parallel texts for terminology research. They are often consulted to get background information or to get a better understanding of the subject of the translation. For example, in the bitcoin.org website, a wealth of information about the principles Bitcoin is based upon and how it functions can be found in the ‘Introduction’ and FAQS sections.

Furthermore, web pages are often consulted in order to check spelling variations or to see whether a term is routinely translated into the target language or whether it is borrowed from the source language. Multilingual websites provide a wealth of terminological resources, for example, the United Nations website\textsuperscript{145} provides content in English, French, Spanish, Russian, Chinese and Arabic.

Multilingual websites are also an excellent source of parallel texts, which can be trawled through to compile DIY corpora. Many of these websites, especially the high-profile ones, are already harvested by linguistic search engines such as Linguee or, indeed, by many existing purpose-built corpora. However, the biggest problem translators have is access to original content written in L1, as opposed to translated content, which is sometimes difficult to identify. Whenever they look for equivalents, translated content is fit for purpose, but when they look for usage in the context of natural language production, they have to seek texts written in that language. Therefore, a range of information seeking skills is needed to be able to identify the web page content suitable for the particular needs of translators.

\textsuperscript{144} http://www.worldwidewebsize.com/

\textsuperscript{145} http://www.un.org/
Online documents

Online documents such as pdfs, PowerPoint presentations or online word documents can often serve as valuable sources of information for translators. Translators are often encouraged to use file extension as a filter. This is because documents in those formats are perceived to be of higher quality as they would have been prepared for publishing or presentation purposes. To the best of my knowledge, there are currently no repositories of online documents such as these for translators; however, there are many websites that offer presentations and other professional content, for example, SlideShare146 which allows users to upload and share presentations, infographics, documents, videos, pdfs, and webinars. Professional content stored in reports, presentations or research papers can also be found on the web pages of well-established companies, organisations, academic institutions and government bodies.

Machine translation

After the crisis following the ALPAC Report nearly fifty years ago (ALPAC, 1966), machine translation has once again, in the last two decades or so, become the main area of development in translation technology, bringing with it much-heated debate about its quality, relevance and usability for professional translators147. According to Van Der Meer & Ruopp (2014), MT is no longer perceived as a “goal on its own”, but as a technology that “empowers many new forms of communications built into wearable technology, search, social media, Internet of Things, Apps” (Van Der Meer & Ruopp, 2014).

The use of machine translation in human translation production has become one of many strands of MT use in general and its adoption by professional translators encompasses only some of its uses. Van Der Meer & Ruopp (2014) report that there are five main uses of machine translation today: gisting, search and discovery, post-editing, speech translation and sentiment analysis. Of these five uses, two are practised by translators: the search and discovery mode and the post-editing mode. In the search and discovery mode, a translator uses MT output for suggestions to individual translation problems. In the post-editing mode, the MT is used to provide the entire target translation which is then revised by a translator.

There is no doubt that the language coverage and quality of machine translation have improved tremendously in recent years. Whilst, in general, it still does not achieve publication-quality performance, “state-of-the-art systems can now deliver a level of quality that makes the post-editing of raw machine output by human translators a viable and cost-effective alternative to translation from scratch” (Germann et al., 2014:iii). In the post-editing mode, AMTA (online) reports proven productivity gains of between 30% and 300% depending on language pair, content type and complexity, experience of the post-editor, domain knowledge, quality requirements, use of automatic QA tools, and quality of training data and reference material. However, to the best of my knowledge, no data exists with regard to productivity gains using MT in the search and discovery mode.

For professional translators, MT is a still a hot topic sparking off heated debates, many of which can be found on discussion fora such as ProZ or Translators Café. A more scientific account of the use of MT by professional translators reveals a rather more balanced perspective which takes into account

146 http://www.slideshare.net/
147 For a concise history of machine translation refer to Hutchins (2007)
the pros and cons of using MT. In the DGT-wide\textsuperscript{148} survey of the MT service at the European Commission based on 763 responses from translators working for DGT, Fontes (2013:11) reports that translators find MT beneficial for the following reasons:
- as a typing aid
- as a source of inspiration for alternative translations available in translation memories
- as a source for a quick draft which can then be improved
- helps cope with heavy workloads
- helps gain time for more thorough research.

The drawbacks, reported by those translators who tried MT but did not use it, included:
- the need for too many changes
- poor quality of MT
- MT disrupted working methods
- MT is not a time-saver as terminology still needs rechecking
- MT gives rise to mistakes translators would otherwise not have made.

In terms of MT performance, the biggest shortcomings appear to be:
- grammar/morphological rules of the target language
- distortion of meaning
- calque of word order from the source text
- terminological inconsistency.

It would, therefore, appear that professional translators, despite occasional prejudice and lack of awareness, can provide good judgment of new technologies and that the rate of adoption of certain technologies by them is a measure of their usefulness. The actual usefulness of MT to a translator depends on many factors, such as the language pair and the linguistic structure of each of the languages, the availability of training data in a given language pair or the quality of the training data. There are four main types of MT which require different approaches and are therefore linked to the quality of the output:
- rule-based (RBMT) - generates an output based on morphological, syntactic and semantic analysis of both source and target languages involved in a translation task
- example-based (EBMT) - generates output based on a corpus of translated examples, serving as a model to the MT system on which to base new translations (Somers & Fernandez Diaz, 2004:1)
- statistical (SMT) - generates output “based on statistical modelling of the word order of the target language and of source-target word equivalences” (Somers, 2003:513)
- Neural MT (NMT) – the most recent breakthrough in machine translation inspired by the workings of the human brain, most notably the function and structure of the brain’s biological neural networks, closely related to areas of machine learning such as pattern recognition, signal processing and artificial intelligence. Major stakeholders in MT such as Google\textsuperscript{149}, Microsoft\textsuperscript{150} or Systran\textsuperscript{151} are already reporting improvements in quality when using NMT.

\textsuperscript{148} DGT (Directorate-General for Translation) is the European Commission’s in-house translation service.

\textsuperscript{149} https://research.googleblog.com/2016/09/a-neural-network-for-machine.html

\textsuperscript{150} https://arc.applause.com/2016/02/19/microsoft-translate-neural-networks/

- hybrid (SMT/rule based/neural) - uses a combination of different approaches to overcome their limitations.

Machine translation systems are available ‘out of the box’, e.g. Google Translate or Microsoft Translator (Bing), or they can be customised to the user’s individual needs by companies specialising in MT, e.g. Pangeanic, Asia Online or TauYou.

Recently, adaptive machine translation in which suggestions are being generated based on translator’s input is making headways in systems such as MateCat, Lilt or SDL Trados Studio. This type of MT is seen as a paradigm shift in the way this technology is used as it puts translator in charge of accepting or ignoring the suggestions. However, despite its many advantages (e.g. speed, decreased need for research) it introduces a new phenomenon - that of ‘priming’, i.e. the influencing translators before they even consider their options (Green et al. 2013); something that is yet to be researched and considered as a new consequence of using technology in the translation process.

To conclude the debate on the main types of resources for translation-related research, all the above-mentioned types of resources constitute separate types, but they differ in terms of their nature, purpose, mode of creation, quality control etc. These differences exist not only between the types but also within each of the types. The challenge lies in providing a classification of these resources that reflects this variety, but also recognises similarities. Therefore, for the purpose of the present work, a way of systematising these resources had to be conceptualised in order to provide a framework for the analysis. In the following sections, firstly, an attempt will be made to outline previous classifications of resources to see which aspects of these classifications could be adopted and, secondly, a more comprehensive classification of online resources according to different criteria will be proposed, based on the previous classifications and proposing new ones. Thirdly, classifications deemed most meaningful for the purpose of this study will be selected.

4.3.2. Previous classifications

Although no formal, up-to-date taxonomies of online resources exist, certain references to how resources can be categorised can be found in translation studies literature. A classification that reflects the resources that were available in the pre-Internet era is provided by Nord (2009:209). In her typology of translation aids she distinguishes between texts, persons and objects, further subdivided into other types of aids. It is interesting to see that all of these resources, plus many new types, are currently available on-line. This includes ‘persons’ as we can now communicate with experts through the Internet and ‘objects’, for example through image search via a search engine.

152 http://www.pangeanic.com/
153 http://www.asiaonline.net/
154 http://www.tauyou.com/
155 At the time of writing, SDL Trados Studio 2017 in which adaptive MT is a new feature has not yet been released.
A more recent categorisation by Hirci (2013:152) offers a basic grouping of 'translation aids' into dictionaries and other reference aids as shown in Figure 4.22. This classification reflects the division of resources in terms of their medium (electronic, online, paper) and in terms of distinguishing between online resources and those contained within the CAT software. However, all online resources are gathered under one umbrella without reflecting the huge diversification within these resources. Also, the division of all resources into dictionaries and 'other' reference aids reflects the dominant position of dictionaries in the past and does not adequately represent the current position of dictionaries amongst other types of resources which have gained prominence in recent years.
Grego (2010:113) offers a classification of resources based on ‘choice’ and arranges the resources according to two criteria: collaborative/non-collaborative and existing/new as shown in Table 4.1 below.

<table>
<thead>
<tr>
<th>Choice</th>
<th>Non-collaborative / self / individual</th>
<th>Collaborative / others / joint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing / retrieved / deduced</td>
<td>dictionary</td>
<td>forum, mailing lists, field experts, training</td>
</tr>
<tr>
<td></td>
<td>glossaries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>encyclopedias</td>
<td></td>
</tr>
<tr>
<td></td>
<td>customer instructions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>international standards</td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>New / created / induced</td>
<td>web queries</td>
<td>forum</td>
</tr>
<tr>
<td></td>
<td>corpus queries</td>
<td>mailing lists</td>
</tr>
<tr>
<td></td>
<td>own glossaries</td>
<td>wiki</td>
</tr>
<tr>
<td></td>
<td>own translation memories</td>
<td>social networks</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Deductive classification of translation resources according to ‘choice’: a suggested categorisation

Table 4.1 DEDUCTIVE CLASSIFICATION OF TRANSLATION RESOURCES ACCORDING TO ‘CHOICE’: A SUGGESTED CATEGORISATION
Source: Grego, 2010:113

This classification offers an interesting insight into how we can look at online resources based on two important, recent developments: the possibility of web-based collaboration and the creation of own resources. However, the problem with this classification is that the changing technological landscape and new trends make all online resources more conducive to collaborative initiatives. For example, dictionaries now welcome users to submit their own suggestions. MacMillan Dictionary has an open, crowdsourced part of the dictionary where users are encouraged to submit new words for inclusion (see Figure 4.23). Since 2009, 3000 new words were added to the open dictionary and approximately 1500 have become full entries in the Macmillan dictionary (Macmillan Dictionary, 2015:online).

Figure 4.23 MACMILLAN DICTIONARY’S OPEN DICTIONARY SECTION
Similarly, the success of Wikipedia and Scholarpedia bears witness to the new collaborative trends being embraced in the domain of knowledge-based resources such as encyclopaedias. A more diverse perspective is offered by Désilets et al. (2009:np). Here, the resources are categorised based on five different criteria, as shown in Table 4.2.

<table>
<thead>
<tr>
<th>Category</th>
<th>Sub-category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature</td>
<td>Termino-lexicographic</td>
</tr>
<tr>
<td></td>
<td>Corpus</td>
</tr>
<tr>
<td>Languages</td>
<td>Multilingual</td>
</tr>
<tr>
<td></td>
<td>Unilingual</td>
</tr>
<tr>
<td>Availability</td>
<td>Public</td>
</tr>
<tr>
<td></td>
<td>Private</td>
</tr>
<tr>
<td>Specialization</td>
<td>Multidomain</td>
</tr>
<tr>
<td></td>
<td>Single domain</td>
</tr>
<tr>
<td>Quality control</td>
<td>Tight</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Open</td>
</tr>
</tbody>
</table>

Table 4.2 LIST OF RESOURCE TYPES AND THEIR USES. Source: Désilets et al., 2009

These criteria accurately capture the many angles from which resources can be viewed, however, other possible classifications such as the collaborative/non-collaborative aspect, which were not part of the analysis of Désilets et al.’s study, are not represented.

Therefore, since none of the three classifications represent the variety of different features of resources in terms of the many different dimensions and recent developments, in the next section an attempt will be made to enrich these categories and provide a more comprehensive overview of the possible categorisations of resources for translators.

4.3.3. Possible classifications of resources for translators

In order to portray a comprehensive landscape of ‘types’ of resources for translators, different ways of classifying the resources according to various dimensions reflecting recent developments will be considered (see Table 4.3 below). This classification offers a multi-dimensional portrayal of the many resources for translators, although the fuzzy boundaries between them mean that it is not always possible to fit a resource neatly into a box.

<table>
<thead>
<tr>
<th>Classification according to:</th>
<th>Termino-lexicographic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Nature of resource</td>
<td>Text-based</td>
</tr>
<tr>
<td>2. Purpose of resource</td>
<td>Linguistic</td>
</tr>
<tr>
<td>3. The number of supported</td>
<td>Extra-linguistic</td>
</tr>
<tr>
<td>languages</td>
<td>Multilingual</td>
</tr>
<tr>
<td>4. Specialisation</td>
<td>Bilingual</td>
</tr>
<tr>
<td></td>
<td>Unilingual</td>
</tr>
<tr>
<td></td>
<td>Specialised</td>
</tr>
<tr>
<td></td>
<td>General</td>
</tr>
</tbody>
</table>
### According to the nature of resource

When considering resources from the point of view of the nature of the data they contain, they can be divided into **termino-lexicographic** and **text-based** resources. A similar distinction was drawn by Désilets et al. (2009), who distinguish between termino-lexicographic resources and corpus-based resources, where termino-lexicographic resources include dictionaries, terminology databases and lexicons while corpus-based resources can include “any body of texts, including Translation Memories and bitexts, as well as documents related to the source texts and Websites” (Désilets et al., 2009:np). In the present work, the corpus-based resources referred to by Désilets et al. are called **text-based** resources. This distinction can be expanded upon further by considering termino-lexicographic resources to be those clustered around a single word, word-compound or expression, and text-based resources to be based on a larger chunk of text.

Termino-lexicographic resources encompass both lexicographical and terminological resources. Wartburton (2010:np) explains that lexicographical resources such as dictionaries, glossaries, word lists, and MT lexicons are lemma oriented, deal with general purpose language and encompass all (or the main) meanings of a word. Terminological resources such as term banks, termbases, thesauri, lexical ontologies, controlled authoring dictionaries or search engine resources are mostly concept oriented, deal with special purpose language and refer to one sense of the word that is being researched (Wartburton, 2010:np)\(^\text{157}\). However, “given the fact that lexicographic resources and terminological resources have things in common and potential points of synergy, and that many so-called terminological resources are lexicographical in structure” (Wartburton, 2010:np) there seems to be a trend towards drawing these two types of resources together (Wartburton, 2010:np; Sager, 2001:761). For the purpose of general classification, the differences

\(^{156}\) Although in literature, online resources are often called electronic, in this study a distinction is made between resources available online and those which are not. Electronic resources are those digital resources that are not available online and therefore constitute a separate category.

\(^{157}\) The differences between lexicographical and terminological resources are more complex; however, for the purpose of this overview, a simplified distinction will suffice.
between terminological and lexicographic resources are not of paramount importance, hence the use of the term 'termino-lexicographic resources'.

In contrast to termino-lexicographic resources, which revolve around a single word/expression or concept, text-based resources revolve around a text and include not only purpose-built corpora but also any collections of texts such as books, websites, translation memories, knowledge-based resources such as Wikipedia, online documents and so on.

**According to the purpose of resource**

The nature of translators’ work is that they seek information of such a varying nature that a large ensemble of resources is often needed to satisfy their research needs. This means that the resources they consult need not necessarily be purpose-built for language or translation uses per se. On the contrary, as 15 petabytes (15 million gigabytes) of data is generated every day on the Internet and is expected to increase twenty-fold by 2020 (Wiggins, 2012), translators are taking advantage of this information, especially as web searching becomes more sophisticated. Therefore, resources for translation-related research can be divided into **linguistic** (purpose-built for language or translation purposes) and **extra-linguistic** ones (those which were not conceived with language or translation activities in mind).

Linguistic resources encompass all resources built for language learners, language researchers or translators. These include resources such as dictionaries, thesauri, ontologies, corpora, or linguistic search engines such as Linguee or WebCorp. Extra-linguistic resources, on the other hand, encompass all other material that translators might want to consult that was not specifically produced for language purposes. These include resources such as knowledge-based resources (e.g. Wikipedia or other encyclopaedias), websites, online documents or generic search engines such as Google.

**According to the number of supported languages**

From the point of view of the number of languages a particular resource supports, Désilets et al. (2009:np) divide resources into unilingual (monolingual) and multilingual. In the present classification, a bilingual category is added to distinguish the resources that are concerned only with two languages such as bilingual dictionaries, glossaries or termbases from those that support more than two languages. Although in the online reality, termino-lexicographic bilingual resources are increasingly being expanded to include other languages, there are still many bilingual resources in existence, especially for the lesser-known languages, for example, Kateglo\(^{158}\) for Bahasa Indonesia, the official Indonesian language. Also, the bilingual category includes the many electronic bilingual dictionaries that are still much in use as well as those resources that are by design bilingual, such as translation memories or bitexts.

Therefore, according to the number of supported languages, it is proposed that the translation resources are classified as: **monolingual**, i.e. those that support only one language (e.g. encyclopaedias, monolingual dictionaries, parallel texts, monolingual corpora etc.); **bilingual**, i.e. those that support two languages (e.g. bilingual dictionaries, glossaries, translation memories, bitexts etc.); and **multilingual**, i.e. resources that support three or more languages (e.g. wikis, search engines, multilingual dictionaries, multilingual ontologies, CLIR resources, machine translation etc.).

\(^{158}\) http://kateglo.com/
According to specialisation

Resources used by translators can also be divided into general and specialised. The general resources encompass any resources that do not deal specifically with any particular domain or subject but cover a variety of areas. These can also be called ‘multi-domain’ and encompass resources such as OED, Wikipedia, TAUS Data, Google Translate, IATE termbase, WordReference Forum159 or the TenTen Corpus Family (Jakubićek et al., 2013:np). Specialised resources encompass any resources that are single-domain and address one specific area, subject or a specific aspect of an area or a subject, for example The Dinosaur Anatomical Dictionary160, Glossary of Environmental terms161, a customised MT engine for the wind power industry, a website dedicated to Bitcoin162 or the Microsoft Terminology portal163.

According to the QA of the resource

A distinction taking into account the quality assurance process involved in the creation and maintenance of a resource was proposed by Désilets et al. (2009:np) to determine which resources were preferred by translators, depending on the level of authority of the creators of the resource. Three ‘gradients’ of quality assurance - tightly controlled, moderately controlled and open resources - were used to determine this. In order to avoid the confusion between what Désilets et al. (2009:np) call ‘open resources’ with ‘open source’, their classification can be adapted as follows: tight resources, such as IATE, OED or Termium, are those resources that are carefully crafted and revised by linguists, terminologists or revisers; moderate resources, such as bilingual sites produced by reputable organisations, internal termbases or translation memories, which are produced by people working in reputable organisations and have therefore a degree of quality assurance (they could also encompass resources with mixed content, such as Linguee, consisting of an editorial dictionary, which can be said to be tightly controlled, and a linguistic search engine, which is not); relaxed resources, such as Wikipedia, publicly available glossaries, discussion fora or any websites available on the Internet, are those resources whose content is not monitored by language specialists.

Whilst it is relatively straightforward to make these distinctions in theory, in practice this might prove more difficult as it is not always possible to determine what quality assurance procedures exist for a given resource, especially for an extra-linguistic one.

According to collaborative functions

The level of permitted collaboration within a resource is a distinction that could only be made relatively recently. This is due to the developments in the architecture of the web, from the static, read-only nature of web 1.0, to the dynamic, read-write web 2.0, where users can modify the web content. The resources which are hosted on static, read-only websites belong to the group of closed resources, whereas those hosted on dynamic, read-write websites belong to the group of open164 resources.

159 http://xen.wordreference.com/
160 http://www.bestonlinecollege.org/dinosaur-anatomical-dictionary/
161 http://www.nrdc.org/reference/default.asp/
162 https://bitcoin.org/en/
164 Not to be confused with ‘open source’
Web 2.0 architectures started permeating the Internet in 2000, and gradually websites have become increasingly dynamic, with collaborative, user-generated content, interactive discussion fora and, more importantly for translators, with the possibility of providing feedback. This brought a shift in the perception of authority with regard to assessing the quality of content. Whereas before, information found on the Internet could not be openly discussed or challenged, now, the possibilities for asking questions and offering alternatives has shifted the authority from the ‘unchallenged specialist’ to peers. One of the best examples is ProZ, where users not only contribute their glossaries but also suggest possible solutions in response to peers’ questions and build their reputation by obtaining KudoZ points. Another example is TermWiki[^165^], where users showcase their profiles by defining or translating terms. It is, therefore, important to observe translators’ behaviour in this respect in future research to see whether the shift from the authoritative nature of web 1.0 to the peer-based nature of web 2.0 has influenced choices and the perception of quality of the content.

Another interesting aspect pertinent to this classification is how technology development affects the amount of ‘freedom’ of choice translators have. Grego (2010:110-111) argues that the Internet has afforded translators more choice but less freedom as “before the choice of how to relate a text to its context was only the translator’s […] today s/he is surrounded by and immersed in - off and online - great numbers of theoretical and practical resources supporting him/her”. Grego also argues that the research process is subjected to a lengthier decision-making procedure due to the multitude of choices (ibid.). It could be argued, however, that the very nature of peer-based resources actually affords greater freedom. Translators now have resources at hand from which they can see how many other translators dealt with the same problem. Being able to see that many possibilities can be acceptable, even in the same context, could be seen as liberating for translators who can now accept or refute a solution and assume full responsibility for their own choices. For freelance translators this was more difficult to achieve before the Internet or even in an earlier, read-only, authoritative Internet, unless they consulted several colleagues on the phone.

*According to the accessibility of resource*

The accessibility of resources refers to whether a resource is paid for by the translator or free of charge. Before the Internet became mainstream and free resources became available, translators used to spend substantial amounts of money on building repositories of reference material and paying subscriptions to dictionaries and other reference sources. As one translator on the ProZ forum says, “in order to start translating, I had to invest about $6,000.00 in paper dictionaries […] There even used to be a company that would rent bilingual technical dictionaries to translators if and when you had a specific project - there was an option to purchase at the end of the rental period […] and a new computer cost around $1500.00” (ProZ, 2015b). Today, there are still subscription-based resources in existence, such as the Oxford English Dictionary; however, most of the online resources can now be enjoyed free of charge.

The roots of this development can be traced back to the advent of the open source movement in 1998, which coincided with the development of the web from web 1.0 to web 2.0. In this new approach to software development, users were encouraged to become developers and thus create software that offered an alternative to expensive, proprietary offerings. This movement impacted not only the pricing

[^165^]: http://en.termwiki.com/
models of proprietary software but also brought about a change in software development by implementing user feedback. As users could no longer be locked into expensive, inflexible systems but could take their own initiative in the development of their own solutions that arose in response to their own needs, open-source solutions began to offer viable alternatives to proprietary ones. Nowadays there is an open-source or a free to use version of almost any kind of software available to users should they choose to use it.

Online resources followed suit and publishers now tend to provide their content for free whilst looking for alternative ways of generating revenue. EU and other regional initiatives for resource creation, inexpensive and user-friendly technology enabling data gathering and content creation, together with new models of collaborative practices and crowdsourcing have also contributed to the fact that translators no longer have to invest heavily in their library of resources, but can access high quality sources from their computer or a mobile device free of charge.

According to the availability of resource

Désilets et al. make a distinction between private and public resources, private resources being those “which could only be accessed by certain translators, typically those working for a particular client or translation office” (Désilets et al., 2009:np) and public being those available to anyone, including the fee-based ones. Therefore, private resources encompass private glossaries and any paper or electronic resources owned by a translator as well as internal termbases, TMs, style sheets or anything made available to a translator by a third party. Public resources include the content of the web in its entirety.

According to the medium of resource

The medium in which a resource is contained can be paper, electronic and online. Paper resources are all printed reference material, including dictionaries, lexicons, encyclopaedias and other printed matter of use to translators. Electronic resources are those stored on CD-ROMs, DVDs, memory cards and other mobile media that can be read by a computer or other electronic device. Online resources encompass all reference material that can be accessed through the Internet.

As with paid resources, paper resources are being largely replaced by online resources (Rundell, 2013) and electronic resources, although still in use, are nowadays used sporadically (see section 5.2.2.). This of course very much depends on language and Internet access in any particular part of the world.

According to the data storage methods

Resources can also be classified according to the way they are stored, i.e. locally, on a server or in the cloud. Local storage refers to any data that can be accessed without an Internet connection. This includes paper resources, material stored on CD-ROMs, DVDs, memory cards as well as any data stored directly on a hard drive, such as personal glossaries or translation memories that are not cloud-based etc. Server-based resources refer to any websites that are not cloud-based and whose data is all stored on a local server, for example, IATE. Cloud-based resources are those that use distributed solutions for their storage, i.e. data can be found on a server anywhere in the world. These encompass any search engines, linguistic or not, translation memory repositories such as TAUS Data or translation memories that are accessed in the cloud.
such as MATE-CAT\textsuperscript{166}. Although technology is moving towards the cloud-based model which essentially enables data sharing, there are many security issues that prevent translators from fully embracing the new model. Another issue to consider is the lack of Internet connectivity, which can prevent users from accessing Internet resources. Therefore, all three types of storage are still used by translators today.

4.3.4. The classification of resources in the present study

As the present study examines the use of online resources by professional translators, certain decisions about the angles of data analysis had to be made. Given the scope of this study, analysing all data against all ten categories mentioned in section 3.1.2. would have been not only impossible but also counter-productive. Not all these categories provide the insight necessary in order to further arguments in this thesis and whilst some are useful in portraying a general picture of the preferences and habits of professional translators, they might not necessarily offer much insight into the portrayal of individual resource behaviour. For example, to see whether paper resources are indeed in decline, the medium of the resource was considered in section 5.2.2. and, therefore, the three types of resources, online, electronic and paper were taken into account. However, since it was established that paper resources are indeed in decline, it was not deemed insightful to analyse this aspect in Chapters 6 and 7 which discussed individual differences.

Similarly, only the types of resources are discussed in section 5.2.3. in order to get a general overview of the habits and preferences of translation professionals with regard to their use of online resources. These were analysed in two ways. In the Global Survey and in the profile questionnaire in the Main Study, the classification of resources in terms of their type such as dictionaries, corpora, TM databases, termbases and glossaries, were derived ‘\textit{a priori}’ (i.e. on the basis of the researcher’s assumptions). However, after the analysis of the screen recordings in the Main Study, a much richer and diverse list of resources emerged. It could be said, therefore, that this classification was derived ‘\textit{a posteriori}’, i.e. from observing the actual user (translator) behaviour.

Throughout this thesis, resources are often analysed in a dual manner; in terms of the types (i.e. dictionaries, glossaries, term banks etc.) and specific resources (i.e. actual individual resources such as Real Academia Española\textsuperscript{167} dictionary, ProZ Glossaries or IATE).

However, one of the most important distinctions for the purpose of this study was the one drawn between \textbf{termino-lexicographic} and \textbf{text-based} resources. This is because in Chapter 6 this distinction helps to differentiate those translators who prefer resources organised around a word or a term (such as dictionaries) from those who refrain from using dictionaries in favour of parallel texts and other forms of text-based resources which, together with other aspects of translators’ individual online information behaviour, contribute to the establishing of translator resource behaviour.

\footnotesize{\textsuperscript{166} https://www.matecat.com/\textsuperscript{167} http://www.rae.es/}
5. Common tendencies in translators’ interactions with online resources

The purpose of this chapter is to provide a summary of the findings from the Global Survey and the Main Study with regard to the translation-related online research behaviour of translators. The focus of this part of the analysis is on common tendencies rather than individual differences, which will be discussed in subsequent chapters. Here, three aspects of the collected data will be explored, focusing on who the participants were, what resources they used (and how many), and how their translation-related research appeared to the observer from the process point of view.

Section 5.1. focusses on the description of the samples (the who) drawing on self-declared data from the Global Survey and the Main Study and using triangulation of data from the two respective samples to check the consistency of the findings.

Section 5.2. concentrates on the participants’ resource use (the what) by looking at data generated in the Global Survey and the Main Study and analysing this data by means of methodological triangulation, i.e. using data obtained by different methods of data collection, data type triangulation (observed and self-declared) and various comparisons of the same type of data, mainly observed.

Section 5.3. examines the distribution of research activities in the translation process of the Main Study by considering observed data by means of data comparison.

Certain trends that emerge from this analysis will be discussed in more detail from the perspective of individual differences between translators in Chapters 6 and 7.

5.1. Sample Demographics - The Who

This section presents an overview and a comparison of the demographic information of the samples in both the Global Survey and the Main Study. This comparison is carried out not only to indicate who the participants in the present research are but also to provide an element of validity for the smaller sample of the Main Study (N=16), which will be analysed in subsequent chapters in greater detail than the larger sample (N=540) of the Global Survey. However, due to the fact that the Main Study sample is relatively small, the comparisons in this section need to be treated with a degree of caution.

5.1.1. Type and size of samples

Non-probability sampling, where participants are “selected based on certain non-random criteria” (Bhattacherjee, 2012:69) was used for both study components. This means that in the Global Survey, only translation professionals\(^{168}\) were targeted and for the Main Study, only translators with at least five years of experience and translating out of English were selected.

However, due to the different nature of the two study components, different sample sizes and different sampling techniques were used. Whilst the Global Survey is characterised by a wide, but shallow span which captured the self-declared behaviour of a large group of translation professionals,

\(^{168}\) As mentioned in 1.3., translation professionals include a variety of translation-related professions such as translators, interpreters, terminologists, project managers etc.
the Main Study focusses on the actual behaviour of a selected sample of experienced professional translators and was therefore restricted to a small number of cases.

The Global Survey sample was obtained using a technique called ‘viral sampling’ (Plowright, 2011:43) in which professional and social networks were used to collect data. Blaikie (2000:205) refers to this technique as ‘network’ or ‘chain’. The Global Survey sample was not controlled at the entry point and anyone who uses online resources for their translation-related work was invited to participate. However, a degree of control was exercised in the process of data management. Out of 677 responses, 540 cases were selected. Only questionnaires that were complete (or mostly complete) were chosen to be analysed.

The sample for the Main Study was more tightly controlled due to methodological constraints (see section 3.1.). The participants were required to translate out of English into any other language and have at least five years of experience working as a professional translator. In this type of ‘expert sampling’, participants are “chosen in a non-random manner based on their expertise on the phenomenon being studied” (Bhattacherjee, 2012:67).

Although the difference in the size of the two samples is considerable, it can be said that in many respects they share common features, as will be shown in the following sections.

5.1.2. Profession

The participants of the Global Survey represent a wide spectrum of translation professionals, with an overwhelming predominance of translators (90%), followed by interpreters (10%), project managers (7.5%) and other professions such as terminologists, educators/trainers, localizers, translation business owners and language technology specialists. Some participants classified themselves into more than one category, e.g. translator and project manager. As far as the Main Study sample is concerned, 100% of the participants declared themselves to be translators and 56% were also interpreters. Almost one-third of the sample (31%) was involved in delivering translator training, whether in a commercial or an academic setting. This could mean that the research attracted many translators who perceived this study as an important one, with a potential value for translator training.

![Figure 5.1 Profession of Participants from the Global Survey and the Main Study](image)

Figure 5.1 PROFESSION OF PARTICIPANTS FROM THE GLOBAL SURVEY AND THE MAIN STUDY
5.1.3. Geographical location

The majority (80%) of the participants in the Global Survey were from EU countries, followed by North America (7%), non-EU European countries (4%), Asia (4%) and South America (4%). The remaining 1% of respondents came from Africa, Australia and New Zealand.

In the Main Study the sample was geographically distributed as follows: 50% of the participants came from EU countries, followed by 19% from South America, 19% North America and 6% from Asia and other EU countries.

In terms of where the participants were geographically located, it appears that the samples were relatively similar, the main difference being in the higher proportion of participants from the EU in the Global Survey (see Figure 5.2 below).

![Geographical Location](image)

Figure 5.2 GEOGRAPHICAL LOCATION OF PARTICIPANTS FROM THE GLOBAL SURVEY AND THE MAIN STUDY

5.1.4. Age and gender

The age distribution in both samples was not considered normal, i.e. the data points were not distributed symmetrically around the mean point in a bell shape: they were skewed to the left, especially in the case of the Main Study sample (see Figure 5.3). However, this is not unusual in small-scale studies. As Saldanha & O’Brien point out, “in translation studies research, where small samples or case studies are more common than in other areas such as psychology […] data are often not distributed normally” (2014:196).

![Age distribution in Global Survey and Main Study](image)

Figure 5.3 AGE DISTRIBUTION IN THE GLOBAL SURVEY AND THE MAIN STUDY
The gender distribution was almost identical in both the Global Survey and the Main Study: 64% and 69% of women respectively (see Table 5.1).

<table>
<thead>
<tr>
<th></th>
<th>MALE</th>
<th>FEMALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Survey N=540</td>
<td>36%</td>
<td>64%</td>
</tr>
<tr>
<td>Main Study N=16</td>
<td>31%</td>
<td>69%</td>
</tr>
</tbody>
</table>

Table 5.1 AGE DISTRIBUTION IN THE GLOBAL SURVEY AND THE MAIN STUDY

This distribution is also reflected in the membership of the ITI\textsuperscript{169}, which confirmed that a clear majority of their 2750 members are indeed female (see Figure 5.4).

\textsuperscript{169} This data was obtained through personal communication with the ITI administration team on the 13 November 2014.
5.1.5. Experience

With regard to the length of service in the industry\textsuperscript{170}, the two samples bear some resemblance, especially in the higher bands (over eleven years of experience). However, there was a slight dominance of translators with six to ten years of experience in the Main Study (see Table 5.2 below). This relates to the age distribution, where most translators were in their thirties.

<table>
<thead>
<tr>
<th>LENGTH OF EXPERIENCE (in years)</th>
<th>≤2</th>
<th>3-5</th>
<th>6-10</th>
<th>11-20</th>
<th>21-30</th>
<th>30+</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLOBAL SURVEY</td>
<td>13%</td>
<td>18%</td>
<td>20%</td>
<td>27%</td>
<td>16%</td>
<td>6%</td>
</tr>
<tr>
<td>MAIN STUDY</td>
<td>-</td>
<td>6%</td>
<td>44%</td>
<td>25%</td>
<td>13%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Table 5.2 LENGTH OF EXPERIENCE DISTRIBUTION IN THE GLOBAL SURVEY AND MAIN STUDY

The ‘experience’ scale was designed slightly differently in each part of the study. The Global Survey includes a ‘two years or less’ category, designed to capture the entrants to the profession. In the Main Study, this category was excluded as only participants with five years of experience or more were allowed to participate in order to meet the criterion of being ‘experienced’.

5.1.6. Specialisation

The participants in the Global Survey were asked to declare their specialisations by selecting them from a multiple selection list. The most frequently selected specialisations were as listed in Table 5.3 below:

<table>
<thead>
<tr>
<th>General</th>
<th>Business/Economic</th>
<th>Technical/Scientific</th>
<th>Legal</th>
<th>Localisation</th>
<th>Medical</th>
<th>Literary</th>
<th>Transcreation</th>
</tr>
</thead>
<tbody>
<tr>
<td>71%</td>
<td>58%</td>
<td>57%</td>
<td>35%</td>
<td>25%</td>
<td>25%</td>
<td>13%</td>
<td>11%</td>
</tr>
</tbody>
</table>

Table 5.3 SPECIALISATIONS OF THE GLOBAL SURVEY PARTICIPANTS

Other specialisations mentioned included: Financial, Marketing, Environmental, Tourism, Education/Training and IT.

Although the specialisation categories were pre-assigned in the Global Survey, space was also provided to allow participants to indicate other specialisations. In hindsight, it is debatable whether the ‘General’ category should have been included since it may not have been clear what types of

\textsuperscript{170} It is assumed that the number of years of experience equals the number of years of practice although two participants who declared having 5-10 years of experience said that at the time of contributing to the present study they worked part-time. Further three participants said they worked part-time, but they declared between 11 and 30 years of experience and some said they had worked full time prior to taking part in the present study. Additionally, five participants in the Main Study were involved in translator training which would mean that translating was not their full-time occupation.
translation it covered. This doubt was reflected in this category not being as frequently indicated in the self-categorisation of participants in the Main Study (see Table 5.4 below), in which the participants were able to freely indicate their specialisations (more than one if applicable) without any pre-assigned categories. The most represented categories of specialisation in the Main Study were:

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Business/Economic</td>
<td>19%</td>
</tr>
<tr>
<td>Technical/Scientific</td>
<td>38%</td>
</tr>
<tr>
<td>Legal</td>
<td>63%</td>
</tr>
<tr>
<td>Medical/Healthcare</td>
<td>31%</td>
</tr>
<tr>
<td>Literary</td>
<td>50%</td>
</tr>
<tr>
<td>Marketing</td>
<td>19%</td>
</tr>
<tr>
<td>IT</td>
<td>19%</td>
</tr>
<tr>
<td>Religion</td>
<td>25%</td>
</tr>
</tbody>
</table>

Table 5.4 SPECIALISATIONS OF THE MAIN STUDY PARTICIPANTS

The self-categorisation resulted in the ‘General’ category being indicated by only three participants, which amounted to 19% of the sample, whereas in the Global Survey this category was marked by 386 respondents, which amounted to 72% of the sample. The areas of Technical/Scientific, Legal and Literary translation were practiced by a similar proportion of translators in both samples (see Figure 5.6); however, other areas such as Localisation, IT or Religion showed greater differences. There was a 20% difference in the proportion of translators specialising in Business/Economic/Financial translation to which the source text for the Main Study belonged as 58% of the Global Survey sample and 38% of the Main Study sample declared specialising in these fields.

As illustrated above, the specialisation category showed the greatest differences between the two samples. A methodological issue of categorisation needs to be highlighted here as it would appear that that the ‘General’ category is not intuitively associated with a specialisation, but rather the lack of it.
According to Martin (2011: online) the word ‘specialisation’ has become fuzzy and essentially remains largely unspecified. He argues that

[for the concept of specialization to be of any real use within the translation industry subject areas of specialization will have to be more specific than such broad categories as ‘legal,’ ‘business’ or ‘technical,’ which do not describe the types of documents a given translator is capable of translating in sufficient detail. Such categories will have to be broken down into relevant sub-categories that reflect specific types of knowledge and skills while also constituting relevant and viable areas of specialization]

(Martin, 2011:online)

Martin calls for a “recognized taxonomy of translation categories” (Martin, 2011, online) which would be helpful not only for translators to market themselves more accurately but also for the purposes of research in translation studies.

The specialisation in Business/Economic/Financial translation is an important variable for the data analysis as the text in the translation task of the Main Study spans these domains. Despite the low percentage of participants in the Main Study declaring specialising in these domains, as many as 81% of them confirmed that even though they may not specialise in this type of translation, they felt comfortable with the text that was chosen for translation (see Figure 5.7 and Figure 5.8).

---

**Figure 5.7 PROPORTION OF MAIN STUDY PARTICIPANTS SPECIALISING IN BUSINESS/ECONOMIC/FINANCIAL DOMAIN**

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>44%</td>
<td>56%</td>
</tr>
</tbody>
</table>

**Figure 5.8 PROPORTION OF MAIN STUDY PARTICIPANTS COMFORTABLE WITH THE SOURCE TEXT WHICH BELONGED TO BUSINESS/ECONOMIC/FINANCIAL DOMAIN**

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>81%</td>
<td>19%</td>
</tr>
</tbody>
</table>
5.2. The use of online resources - The What

In this section, the focus will be on the use of resources by the participants in this study, covering the self-declared data from the Global Survey as well as the self-declared and the observed data from the screen recording part of the Main Study. The aim is to portray the common tendencies of the participants in translation-oriented research behaviour and to provide some insight into the use of online resources by professional translators to identify similarities, differences and ranges of behaviour. In order to do that, this use of resources will be explored in different ways. The following are the areas that will be explored in the subsequent sections:

- the time spent on translation-related research
- the medium of the resource (online vs. paper)
- the type of resource (e.g. dictionaries, glossaries, search engines etc.)
- the specific resources (e.g. WordReference dictionary, ProZ Glossaries, Google)
- the frequency of use of the types and the specific resources.

While observing the general patterns of research behaviour, the individual differences between the translators observed in the Main Study will be briefly highlighted as signposts to the discussion in Chapters 6 and 7, where these individual differences will be explored in greater detail.

5.2.1. Time spent on translation-related research activities

The time spent on translation and the different phases of the translation process have been addressed in process-oriented translation research on many occasions. For example, Englund Dimitrova (2005:86-88) distinguished three phases: pre-writing, writing and post-writing, which were measured based on video recordings. In the context of dictionary use in foreign language learning, Tono (2001:112) timed her participating students in order to analyse the translation time differences between the more proficient and the less proficient learners.

There are also numerous references to the time translators purportedly spend on research; however, none of them provides an indication of the sources from which their data emerged. For example, Mayorcas-Cohen† (1986:75) reported around three decades ago that “surveys have found that translators may spend 25-60 percent of their time on [terminology] research”. She did not, however, make references to any particular surveys. Liddicoat & Baldauf (2008:31) comment on Kerpan’s (1991) study in which she suggests that translators spend 45% of their time on translation research, but again, the empirical basis for this figure is hard to establish. A more recent and much more comprehensively documented account by Champagne (2004:30) based on twelve case studies and two focus groups revealed that “terminology research is required for 4% to 6% of all words in a text” and that “experienced translators spend about 20% to 35% of their work on terminology activities” (Champagne, 2004:30). For a new translator, Champagne declares, this percentage might be as high as 40% to 60%. This, however, includes all terminology activities, including creating terminology entries in a termbase, which, Champagne reports, takes about fifteen minutes per entry (Champagne, 2004:30). More recently, and from a technology developer’s point of view, Salzberg (2009:online) reports that, according to Dr Takeshi Abekawa, chief developer of the Qreddit

† Mayorcas-Cohen is a professional translator; however, she refers to surveys carried out by others, although no particular studies or names are mentioned.
system, “translators spend 20 to 60 percent of their time consulting online dictionaries or looking up information on the web, which blocks the flow of translation” (Salzberg, 2009:online). Again, no empirical basis for these statements is provided. The most recent reference to the time spent on terminology comes from the 2014 Terminology Summit in Barcelona, where it was reported by Tatiana Gornostay of Tilde that “technical writers and translators spend up to one-third of their time on terminology work and, in some cases, it can consume an even greater share of working time” (Gornostay, 2014:online).

Although the above accounts provide estimates of how much of their time translators spend on research activities, mostly related to what is considered to be ‘terminological research’ none of them has been empirically verified. To the best of my knowledge, the part of the translation process devoted to the interaction with external resources has not hitherto been empirically measured. The only study that makes an attempt at measuring one translator’s interactions with the texts (translation time) and pauses (presumably to deal with translation problems) is a study carried out over a five-week period before 2000 and reported on in 2004 by Hébert-Malloch. However, inaccurate calculations, as mentioned in section 2.2., prevent any conclusions from being drawn. Also, PACTE (2009) reportedly obtained data on time spent on searches in their study of thirty five professional translators and twenty four teachers; however, it is not clear where these findings are reported. The present study is attempting to fill this gap by separating what can be regarded as translation activity from research activity through observed data, providing an empirically based indication of the time spent on interactions with resources as opposed to interactions with the source and target texts. Research activities are deemed to be those activities that happen outside the text production window, but are limited to only those activities that are directly linked to researching for translation. The conceptualisation of the translation process in this study as a series of interactions with texts and resources facilitated such measurement, although the methodological difficulties of separating the two types of interactions should not be underestimated (see section 3.4.). The figures obtained through the objective, empirical measurement carried out in the translation task of the Main Study are complemented by the subjective, self-declared data obtained from the profile questionnaires in the Main Study as well as from the extensive Global Survey.

According to data obtained in the Global Survey, translators spend on average 21% of their translating time on translation-related research in familiar domains and 42% in unfamiliar domains. In the Main Study, it was observed that average research time was 30% in familiar domains and 36% in unfamiliar domains. Interestingly, in familiar domains, the times declared by the participants in the profile questionnaires of the Main Study were exactly the same as those observed in the translation task. In unfamiliar domains, the respondents exceeded the observed Main Study data by 14%, which could be explained by the unpredictability factor of what constitutes an ‘unfamiliar’ domain. Figure 5.9 below provides a comparison of the Global Survey data with the Main Study data and shows how the time spent on research activity varied between the familiar and the unfamiliar domains and between the two study components.
Without the distinction between familiar and unfamiliar domains, the self-declared average time spent on research activities by Main Study participants amounted to 40% and the observed average time to 33%, which is lower than the 45% reported by Kerpan in 1991 but closer to the “one-third of [translators’] time” suggested over twenty years later, in 2014, by Gornostay (2014:10).

However, whilst it is useful to observe the group averages for the time spent on translators’ research for the purpose of comparison with the figures reported by others, it is also interesting to look at the individual differences. As already suggested by Mayorcas-Cohen (1986:online) and in Salzberg’s article (2009:online), translators do vary considerably in the amount of time they devote to research, spending between 20% and 60% of their translation time on interactions with external resources. Such variations are confirmed in the present study. For example, the Global Survey data shows that, in unfamiliar domains, some translators reported spending as little as 20% or as much as 70% of their time on research. In the translation task of the Main Study translators were observed to spend as little as 0% and as much as 56% of their time on research (see Figure 5.10). Rather surprisingly, both the lowest and the highest figures were recorded for translators who were familiar with the subject or specialising in Business/Economic/Financial translation to which the source text belonged. Figure 5.10 below shows the percentage of research time observed for each Main Study participant. The dark blue columns indicate those participants who do not specialise in the Business/Economic/Financial domains.¹⁷²

¹⁷² As noted, not specialising in the particular domain did not preclude participants from feeling at ease with the text.
Translators who do not specialise in the source text domain spent between 23% and 49% of their time on research, with an average of 35.5%. Those who do claim to specialise in the business/economic/financial domains spent between 0% and 56% of their time, with an average of 29.8%. This would suggest that there is no clear connection between declared domain expertise and the actual time spent on research, which would point to the fact that the time spent on research could be linked to the research style of a translator (see Chapter 7).

Last but not least, based on 540 respondents’ self-declared data in the Global Survey, time spent on research activities reportedly takes a similar amount of time regardless of whether translators use technologies such as translation memory (TM) systems or terminology management systems (TMS) (see Figure 5.11). In the case of non-customised MT, data shows that in non-familiar domains, translators seemed to spend slightly more time on research activities when using non-customised MT such as Google Translate. This would point to the fact that the use of technologies such as translation memory or terminology management does not necessarily shorten the research time, or last least such is the THE SELF-DECLARED USE OF DICTIONARIES IN THE GLOBAL SURVEY WAS LOWER THAN IN THE. This, perhaps could be explained by the fact that although more volume of text can be translated with these technologies it is likely that at the same time, more potential problems requiring research might arise precisely due to this increased volume of translated material.

![Observed % of time spent on research, Main Study data](image)
5.2.2. The medium - online vs. paper resources

This study’s focus on online resources reflects the trend in lexicography of moving away from printed resources to those available online. The growing tendency to rely on online resources reverberates throughout various publications. Even as early as 2001, Wright & Budin (2001:845, original emphasis) stated that “the Internet has become THE primary information resource for language professionals”. Hirci (2013:151) goes further by saying that “[w]hat used to be a common everyday translation practice in the past, is now […] considered highly outdated and time-consuming, if not completely inefficient. Working with tons of paper dictionaries and encyclopaedias are a thing of the past”. Many reference resources such as Encyclopaedia Britannica have recently decided to phase out their printed versions. As of 2013, Macmillan Dictionaries has become available exclusively online and the publishers of the Oxford English Dictionary “have also implied that its third edition, […] is unlikely to appear in print” (The Guardian, 2012:online). The chief editor of Macmillan, Michael Rundell, says that “[t]he traditional book format is very limiting for any kind of reference work […]. Books are out of date as soon as they're printed, and the space constraints they impose often compromise our goals of clarity and completeness. There is so much more we can do for our users in digital media” (The Guardian, 2012:online). For Rundell, despite the criticisms of limited ‘browseability’ of online dictionaries which facilitates serendipitous discoveries as well as the connectivity problems associated with their use, the direction towards online resources is clear. The new generations of users, he says, “are digital natives […] and the web will always be the first (and, usually, only) port of call” (Rundell, 2013:5-6).
To balance the discussion, it must be said that there is empirical evidence that paper and electronic resources are indeed both still in use. For example, Massey & Ehrensberger-Dow report that a small number of translators in their study reported using paper and electronic resources, although, much to their surprise, some of them did not appear to know the difference between the online and electronic resources such as CD-ROM or DVD (Massey & Ehrensberger-Dow, 2011b:204). It is also worth considering the voices of professional translators who say that although they might be in favour of technological advancements, some feel that by moving away from printed reference material, “valuable thinking time has been stolen from [them]” (Davies, 2015:44). “Some of these time-saving technologies have not, in fact, given us more time to think about the difficult bits; they’ve simply shortened the time available to do the job as a whole”, says Davies, echoing other professional colleagues. He ponders “why would we ever entirely give up flicking for clicking?” (Davies, 2015:44).

However, despite criticisms of the online medium and occasional “touching display[s] of bibliophilia” (Rundell, 2013:5), users are indeed embracing the migration of reference material from print to online media, as confirmed by the results of this study. In the Global Survey respondents were asked in what proportions they use paper, online and other resources, such as electronic dictionaries, by assigning a percentage to the three types. Figure 5.12 below shows the question as it appeared in the Global Survey.

![Figure 5.12 QUESTION 8 OF THE GLOBAL SURVEY (THE USE OF ONLINE, PAPER AND OTHER RESOURCES)](image)

The respondents were able to assign a different percentage to each of the types of resources. As an example, while Respondent 99 (R99) declared spending 10% of time on consultations in paper resources and 90% on online resources, R97 spent 90% on online and 10% on other resources, such as CD-ROM (See Figure 5.13).

![Figure 5.13 EXAMPLES OF GLOBAL SURVEY RESPONDENTS' INDICATION OF TIME SPENT ON ONLINE, PAPER & OTHER RESOURCES](image)
Figure 5.14 illustrates the distribution of all the answers, with a clear indication that the majority of respondents used predominantly online resources with a very small percentage declaring the use of predominantly paper and other resources such as CD-ROM.

Figure 5.14 SELF-DECLARED PERCENTAGE OF USE OF VARIOUS TYPES OF RESOURCES ACCORDING TO MEDIUM (ONLINE, PAPER & OTHER) - GLOBAL SURVEY DATA

In the Main Study, the question concerning the use of online, paper and other resources was not explicitly asked. Instead, screen recordings were used to observe the actual use of resources. Out of sixteen, only one person used a paper dictionary and three people used electronic dictionaries. One participant said that they would have used their favourite paper dictionary but because this study examined the use of online resources, they resorted to using an online dictionary instead.

In conclusion, it can be said that the results of this study confirm that the online medium is indeed the first choice for translators, as indicated by the 540 participants in the Global Survey and confirmed by the actual behaviour of the sixteen participants in the Main Study. Therefore, this aspect will not form a part of the discussion of individual differences in Chapter 6.

5.2.3. Types of resources

Knowing what types of resources are being used by professionals in the translation profession is important not only for the technology developers but also for the teachers and trainers in educational settings. This knowledge is also important to enable future diachronic studies in this area to see how the use of resources changes over time.

The use of resources according to their type has been studied extensively in both components of the present study, that is, in the Global Survey and the Main Study. As mentioned in section 4.3.4., two classifications of types of resources were used in this study. The a priori classification was formulated in order to ask questions in the Global Survey and the Main Study, and the a posteriori classification, derived from the observed data after the Main Study was completed, in order to capture the types of

173 For an alternative view resulting from a study of Chinese students see Zheng, 2014.
174 Since the number of paper and electronic dictionaries used by participants in the Main Study was marginal, the few instances of their use was included in these calculations as if they were online, in order to study other aspects of dictionary use, for example, whether they were bilingual or monolingual.
resources not included in the *a priori* classification. Therefore, the *a priori* classification will be used to describe the self-declared data and the *a posteriori* classification will be used to describe the observed data.

**The self-declared use of types of resources**

The self-declared data incorporates data from the Global Survey and from the profile questionnaires of the Main Study. The *a priori* classification used in this section concentrates on the following types of resources as the most likely ones to be used by professional translators: dictionaries, glossaries, termbases, corpora and translation memories.

Participants in the Global Survey and in the Main Study were asked what types of resources they use in their daily work and how frequently they use them. As shown in Figure 5.15, in both study components dictionaries were the most widely and frequently used resources, although not by a huge margin. The self-declared use of dictionaries in the Global Survey was lower than in the Main Study. The second most used, self-declared types of resources were TM databases followed by glossaries. Translator’s own resources were favoured to online ones.

![Figure 5.15 SELF-DECLARED USE OF TYPES OF RESOURCES IN THE GLOBAL SURVEY AND IN THE MAIN STUDY](image)

In what follows, the use of dictionaries will be discussed first, followed by an overview of how other resources are reportedly used by translators in their work.

As mentioned above, dictionaries were declared to be the most frequently used type of resource by the participants in the Global Survey and in the Main Study, with the Main Study participants declaring a more frequent overall dictionary usage (see Figure 5.16 below).

---

175 Here only data relating to general bilingual dictionaries is shown although a question about the use of specialised dictionaries was also asked in the survey. In the Main study, the distinction between general and specialised dictionaries was not made and participants only distinguished between monolingual and bilingual dictionaries.
The self-declared use of monolingual vs. bilingual dictionaries was examined in both study components. From the Global Survey data, it would appear that bilingual dictionaries are more widely used than monolingual dictionaries. For the combined categories of ‘Mostly’ and ‘Frequently’, bilingual dictionaries seem to be used almost 1.5 times more often (73%) than monolingual dictionaries (50%). However, the Main Study data shows this difference not only to be less prominent but the monolingual dictionaries being used more frequently than bilingual ones. The combined ‘Mostly’ and ‘Frequently’ category shows a similar to the Global Survey proportion of bilingual dictionaries being used (72%), but the proportion of monolingual dictionaries was higher, at 87% (see Figure 5.17).

Compared to the use of dictionaries, other online resources such as glossaries, termbases, corpora and translation memories seem to attract less high-frequency users. Figure 5.18 below illustrates this\textsuperscript{176}.

\textsuperscript{176} For this calculation, the Global Survey data relating to the self-declared use of ‘own’ and ‘online’ glossaries, termbases and TM was conflated into one set of data, whereas the data from the Main Study reflects the one category of glossaries, termbases and TM that was asked in the questionnaire.
<table>
<thead>
<tr>
<th></th>
<th>SELF-DECLARED DATA</th>
<th></th>
<th>SELF-DECLARED DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GLOBAL SURVEY, N=540</td>
<td></td>
<td>MAIN STUDY, N=16</td>
</tr>
<tr>
<td><strong>GLOSSARIES</strong></td>
<td><img src="image1" alt="Graph" /></td>
<td><strong>GLOSSARIES</strong></td>
<td><img src="image2" alt="Graph" /></td>
</tr>
<tr>
<td>Mostly</td>
<td>14%</td>
<td>Mostly</td>
<td>6%</td>
</tr>
<tr>
<td>Frequently</td>
<td>34%</td>
<td>Sometimes</td>
<td>38%</td>
</tr>
<tr>
<td>Sometimes</td>
<td>32%</td>
<td>Rarely</td>
<td>14%</td>
</tr>
<tr>
<td>Never</td>
<td>14%</td>
<td>Never</td>
<td>6%</td>
</tr>
<tr>
<td><strong>TERMBASES</strong></td>
<td><img src="image3" alt="Graph" /></td>
<td><strong>TERMBASES</strong></td>
<td><img src="image4" alt="Graph" /></td>
</tr>
<tr>
<td>Mostly</td>
<td>13%</td>
<td>Mostly</td>
<td>7%</td>
</tr>
<tr>
<td>Frequently</td>
<td>28%</td>
<td>Sometimes</td>
<td>50%</td>
</tr>
<tr>
<td>Sometimes</td>
<td>29%</td>
<td>Rarely</td>
<td>14%</td>
</tr>
<tr>
<td>Never</td>
<td>17%</td>
<td>Never</td>
<td>14%</td>
</tr>
<tr>
<td><strong>CORPORA</strong></td>
<td><img src="image5" alt="Graph" /></td>
<td><strong>CORPORA</strong></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td>Mostly</td>
<td>5%</td>
<td>Mostly</td>
<td>3%</td>
</tr>
<tr>
<td>Frequently</td>
<td>17%</td>
<td>Sometimes</td>
<td>45%</td>
</tr>
<tr>
<td>Sometimes</td>
<td>24%</td>
<td>Rarely</td>
<td>17%</td>
</tr>
<tr>
<td>Never</td>
<td>23%</td>
<td>Never</td>
<td>17%</td>
</tr>
<tr>
<td><strong>TRANSLATION MEMORY</strong></td>
<td><img src="image7" alt="Graph" /></td>
<td><strong>TRANSLATION MEMORY</strong></td>
<td><img src="image8" alt="Graph" /></td>
</tr>
<tr>
<td>Mostly</td>
<td>9%</td>
<td>Mostly</td>
<td>6%</td>
</tr>
<tr>
<td>Frequently</td>
<td>17%</td>
<td>Sometimes</td>
<td>25%</td>
</tr>
<tr>
<td>Sometimes</td>
<td>22%</td>
<td>Rarely</td>
<td>25%</td>
</tr>
<tr>
<td>Never</td>
<td>19%</td>
<td>Never</td>
<td>19%</td>
</tr>
</tbody>
</table>

Figure 5.18 SELF-DECLARED USE OF ONLINE GLOSSARIES, TERMBASES, CORPORA & TM IN THE GLOBAL SURVEY AND IN THE MAIN STUDY

140
Online glossaries were reported as being regularly\textsuperscript{177} used by 48% of the respondents in the Global Survey and 44% of the participants in the Main Study while online termbases attracted 41% of regular users in the Global Survey and 57% in the Main Study. Online translation memories and corpora are appeared to be the least used. Online TMs appealed to 26% of the 540 Global Survey participants and to the 31% of the Main Study sample, who said they use them mostly or frequently. Corpora were reportedly the least used resource. In the Global Survey, 22% of the respondents said that they used corpora regularly or frequently, although a higher percentage of regular use, 48%, was reported by participants in the Main Study. In the next section, a brief overview of how this self-declared data matched up to the data observed in the screen recordings will be presented.

The observed use of types of resources

The types of resources observed to have been used in the Main Study will be looked at in terms of their quantity and their diversity. The quantity measurement takes into account the number of resource types used in the translation task excluding repetitive use by multiple translators, and the number of types used in the task including repetitive use. The diversity focuses on the range of types of resources used by the participants. While in the previous section the \textit{a priori} classification was used in relation to the self-declared use of resource types, here the \textit{a posteriori} classification is used with regard to the observed use of these types.

In terms of the quantity of resource types, in the Main Study translators were observed to have used seventeen different types, excluding repetitive use by multiple translators. These are listed below.

1. Monolingual dictionaries
2. Bilingual dictionaries
3. Synonyms dictionaries
4. Collocation dictionaries
5. Orthographic dictionaries
6. Idiom dictionaries
7. Glossaries
8. Term banks
9. Thesauri
10. Concordancers
11. Discussion fora
12. Non-customised machine translation engines
13. Search engines
14. Knowledge-based resources
15. Web pages (parallel texts)
16. Online documents (pdf, ppt, digital documents)
17. Ask an expert

In terms of the number of various types of resources used in the translation task of the Main Study including repetitive use by multiple translators, Table 5.5 below shows how many types of resources were accessed by all participants during the translation task. The most frequently used types of

\textsuperscript{177} Regularly here is used to indicate the ‘Mostly’ and ‘Frequently’ categories.
resources are: web pages (n=55), bilingual dictionaries (n=21), knowledge-based resources (mainly Wikipedia) (n=17) and search engines (n=15), accounting for over 66% of all consultations. For example, a vertical reading of Table 5.5 shows that T1 used four different websites, four different bilingual dictionaries, one knowledge-based resource and one search engine. A horizontal reading tells us, for example, how many different websites were used by each individual participant, totalling 55\(^{178}\).

<table>
<thead>
<tr>
<th>Translator</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web pages</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>18</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Bilingual Dictionaries</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Knowledge-based</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Search Engines</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Monolingual Dictionaries</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Concordancers</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Glossaries</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Discussion fora</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Online documents</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Term banks</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Synonym Dictionaries</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Machine Translation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Collocation Dictionaries</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Orthographic Dictionaries</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Idiom Dictionaries</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Thesauri</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

\(^{178}\) This calculation includes the same type being accessed by different translators, for example glossaries were used by six participants, but all of them used ProZ Glossaries and T1 also used her own.
These numbers were calculated by considering how many resources of each type were used by all individual translators in the task, but with each type being counted only once. This calculation does not show the frequency of use of each of the types, i.e. how many times each type was accessed during the task, which is dealt with in section 5.2.5.

The high reliance on web pages as a type of resource noted in the present study was also indicated in a longitudinal, comparative study by Hirci (2013:155-156), where she reports that the main shift in the use of external resources by trainee translators between 2005 and 2012 can be observed in the use of parallel texts, mainly through Google searches. According to Hirci (2013:156), in 2005 60% of the twenty participants in her study claimed to have used online parallel texts, compared to the 100% who used bilingual dictionaries. However, in 2012, all twenty participants in her follow-up study reported using the Internet on a par with resources such as dictionaries, glossaries and encyclopaedias, thus showing a sharp increase in the use of online parallel texts. The present research would suggest, as shown in Table 5.5, that the use of parallel texts (in this study referred to as web pages) has gained even greater prominence.

The second aspect of the investigation into the types of resources used in the Main Study, apart from their quantity, is their diversity. Research conducted by Désilets et al. (2009:np) shows that translators can, on average, have as many as ten specific resources in their tool box, i.e. ten resources being used for a specific task. With regard to the types of resources used, this diversity was also exhibited by the participants in the Main Study who, during the translation task, used between 0 and 12 types of resources (see Figure 5.19 below), with an average of 6.25.

![Figure 5.19 NUMBER OF TYPES OF RESOURCES USED BY INDIVIDUAL PARTICIPANTS IN THE MAIN STUDY](image-url)
Apart from T2 who did not use any resources, there seems to be a range of between four and twelve types, which shows a considerable diversity. This has contributed to the later classification of the participants according to their research diversity (see section 7.1.1.4.). Appendix 6 shows how many types of resources were used by each of the participants, and how many specific resources of that particular type were consulted.

5.2.4. Specific resources

This section focuses on the specific resources used in this study by examining firstly the self-declared data from both study components and secondly, the observed data from the Main Study. With regard to the self-declared data, due to the fact that this research is language independent, no language-specific resources were taken into account when designing the questionnaires. Rather, potentially useful resources were selected based on their free availability and the range of languages covered. These were presented in a list and a 5-point Likert Scale was used to determine the perceived usefulness of each resource.

Self-declared use of specific resources

In this section, two aspects of the self-declared use of specific resources in both study components will be discussed. Firstly, we shall look at what particular resources are reportedly used by translators in their work and secondly, how many specific resources they normally use during a translation task.

A range of resources appeared to be used by both samples, including ProZ, WordReference, Linguee, IATE, Eurotermbank, EUR-Lex179 and Acronym Finder180 (see Figure 5.20 and Figure 5.21). What is interesting, however, is that virtually all of the named resources were found to be useful to some extent, but the number of resources that the respondents had never heard of was surprisingly high. For example, over 60% of the Global Survey respondents had not heard of Eur-Lex or Eurotermbank, and 40% had not heard of IATE181. Given that the overwhelming majority of the respondents in both the Global Survey and the Main Study come from Europe, these figures point to the fact that not only are certain resources for translators underused, but also that they are either not marketed well enough to reach their potential users or are not given sufficient attention during translator training, a point also made by Massey & Ehrensberger-Dow (2011b:20).

179 http://eur-lex.europa.eu/homepage.html/
180 http://www.acronymfinder.com/
181 It is important to remember that the Global Survey was conducted in October 2012 and that these figures might have changed since then.
Even more thought-provoking is the fact that many of the resources participants were asked about in the Global Survey and in the profile questionnaire of the Main Study, seem to be unknown to professional translators, despite the fact that most of these resources cover multiple languages. Figure 5.22 and Figure 5.23 below illustrate the many resources that the respondents did not know about and have never used.
Many of these resources (or tools acting as resources) were specifically developed for language specialists. For example, WebCorp, a linguistic search engine designed to access the web as a corpus which can help translators observe “how particular words and phrases are used, especially words and phrases which are too new or too rare to appear in any dictionary or standard corpus” (WebCorp - Guide: online). 2Lingual has been developed using new technologies such as cross-language information retrieval (CLIR) allowing translators to retrieve information in two languages simultaneously. One resource based on this technology, Linguee, seems to have been widely adopted, but others, such as 2Lingual have not. This is
Despite the fact that 2Lingual is essentially a Google search in two languages, which ought to make more sense for translators, given the fact that fifteen out of sixteen translators used Google in the translation task of the Main Study. For example, one of the most searched terms in the translation task was fiat money. Table 5.6 below shows the breakdown of steps taken to reach a conclusion on how to translate fiat money for all participants, showing that some translators took up to eight steps using some six resources to find a solution. Figure 5.24 shows that a solution can be obtained in one step (or two, if confirmation of the solution is also sought) through 2Lingual, as the snippets in Polish not only show the translation but also describe the concept of ‘fiat money’.

<table>
<thead>
<tr>
<th>Fiat Money</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
<th>Step 5</th>
<th>Step 6</th>
<th>Step 7</th>
<th>Step 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>Google</td>
<td>ProZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>T3</td>
<td>Bilingual dict</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>T4</td>
<td>Google</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>T5</td>
<td>GT</td>
<td>Google</td>
<td>Website</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>T6</td>
<td>IATE</td>
<td>Google</td>
<td>Wiki</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>T7</td>
<td>TermWiki</td>
<td>Google</td>
<td>Wiki</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>T8</td>
<td>Google</td>
<td>ProZ</td>
<td>ProZ</td>
<td>ProZ</td>
<td>Google</td>
<td>Google</td>
<td>Google</td>
<td>7</td>
</tr>
<tr>
<td>T9</td>
<td>Google</td>
<td>Wiki</td>
<td>Wiki</td>
<td>Google</td>
<td>Wiki</td>
<td>Google</td>
<td>Google</td>
<td>6</td>
</tr>
<tr>
<td>T10</td>
<td>Linguee Wiki Investopedia Wiki Wiki Google encyclopaedia Spec dict</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T11</td>
<td>GT</td>
<td>Bi Spec Dict</td>
<td>Bi Spec Dict</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T12</td>
<td>Google</td>
<td>Google</td>
<td>Wiki</td>
<td>Wiki</td>
<td>Google</td>
<td>Wiki</td>
<td>Google</td>
<td>7</td>
</tr>
<tr>
<td>T13</td>
<td>Bi Dict</td>
<td>Website</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>T14</td>
<td>Bi Dict</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>T15</td>
<td>Google</td>
<td>Mono Dict</td>
<td>Mono Dict</td>
<td>Bi Dict</td>
<td>Forum</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>T16</td>
<td>Linguee Bi Dict Forum</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.6 BREAKDOWN OF STEPS TAKEN BY INDIVIDUAL TRANSLATORS TO REACH THEIR CONCLUSION REGARDING THE TRANSLATION OF THE TERM FIAT MONEY

If looking for confirmation of the solution found in the Wikipedia entry, the same resource can be used by reversing the languages. Figure 5.25 below shows the confirmation in two credible resources (a financial information portal and a reputable Polish encyclopaedia). In this case, the research need would have been satisfied in two steps.

---

182 A step, as defined in section 3.2.3., is a single consultation of a specific resource.
Other little-known resources include resources such as OPUS\textsuperscript{183} and Sketch Engine, which is surprising given the fact that a high number of participants in the Main Study declared in the profile questionnaires that they use corpora. WebiText\textsuperscript{184} (which can access predefined corpora and the web as a corpus) and collections of translation memories such as DGT TM, MyMemory or TAUS Search were also barely recognised in either sample.

The second aspect of the self-declared use of specific resources discussed in this section is the number of specific resources translators normally have in their tool box (i.e. actually in use) during a translation task. In the Main Study profile questionnaire, the participants were asked how many resources they normally use while working on a translation task. Out of the sixteen participants, four said that they normally access 1-3 resources, seven declared using on average 4-6 resources and five respondents stated that they usually use 7-9 resources (see Figure 5.26 below).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.26.png}
\caption{CONFIRMING A SOLUTION TO THE TRANSLATION OF THE TERM FIAT MONEY IN ONE STEP USING 2LINGUAL}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.26.png}
\caption{SELF-DECLARED NUMBER OF RESOURCES IN A TOOL BOX IN MAIN STUDY}
\end{figure}

\textsuperscript{183} http://opus.lingfil.uu.se/
\textsuperscript{184} http://www.webitext.com/bin/webitext.cgi/
The average of these figures would give an indication of five resources per tool box. This is half of the average number observed in the screen recordings of the translation task, which followed the profile questionnaire, and in the Désilets et al. study conducted in 2009.

However, it must be noted that translators often think of resources as only those specifically developed for translation purposes and do not take into account extra-linguistic resources such as search engines or parallel texts, thereby lowering the number they self-declare. This might be due to the following reasons:

a. question not being asked precisely enough in the questionnaire
b. the lack of accepted classifications to which translators could refer
c. the lack of a systematic approach to the area of translation resources through translator training.

Bearing in mind the above possibilities, the observed (as opposed to self-declared) number of resources used in the task is considered below.

**Observed use of specific resources**

As in the case of self-declared use of specific resources discussed in the previous section, two aspects of the observed use of specific resources will be discussed in this section. Firstly, we shall look at which specific resources were observed to have been used in the translation task by the sixteen participants in the Main Study. Secondly, the average number of specific resources used in the task by the participating translators will be established.

The patterns of use of specific resources, as opposed to the types, is more difficult to capture due to the great variety of resources accessed by individual participants. A useful distinction between individual vs. common information needs is offered by Enríquez Raído (2014:115), where the individual information needs are those which were observed to have been used by one participant, and common information needs are those shared by at least two participants. In the present study, this distinction is adopted and extended to the use of resources, where *individual resource* use characterises the use by one participant only, whereas *common resource* use refers to those resources accessed by more than one participant.

A total of 99 specific resources in eight languages were accessed by the sixteen participants during the translation task of the source text containing 412 words. A full list of specific resources used in this task can be found in Appendix 7. Due to the language independent nature of the Main Study, most of the specific resources (81 out of 99) were *individual* as translators used resources in eight different languages. The languages represented and the corresponding number of participants translating into that language in the Main Study are presented in Table 5.7.

<table>
<thead>
<tr>
<th>Language</th>
<th>Spanish</th>
<th>Polish</th>
<th>Portuguese</th>
<th>Brazilian Portuguese</th>
<th>French</th>
<th>Dutch</th>
<th>Hungarian</th>
<th>Indonesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants translating into this language</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5.7 LANGUAGES REPRESENTED IN THE MAIN STUDY
Out of the 99 specific resources observed to have been accessed during the translation task, 20 were *common*, i.e. used by at least two participants. Out of these 20, 11 were used by at least three participants. Table 5.8 below shows these 11 common resources consulted by at least three participants and the number of translators who used a particular common resource.

<table>
<thead>
<tr>
<th>Number of translators</th>
<th>Common resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Google</td>
</tr>
<tr>
<td>2</td>
<td>Wikipedia</td>
</tr>
<tr>
<td>3</td>
<td>Linguee</td>
</tr>
<tr>
<td>4</td>
<td>Bitcoin.org website</td>
</tr>
<tr>
<td>5</td>
<td>ProZ</td>
</tr>
<tr>
<td>6</td>
<td>WordReference</td>
</tr>
<tr>
<td>7</td>
<td>ProZ discussion forum</td>
</tr>
<tr>
<td>8</td>
<td>Google Translate</td>
</tr>
<tr>
<td>9</td>
<td>WordReference discussion forum</td>
</tr>
<tr>
<td>10</td>
<td>IATE</td>
</tr>
<tr>
<td>11</td>
<td>Mt.Gox website</td>
</tr>
</tbody>
</table>

Table 5.8 FREQUENCY OF USE OF THE MOST USED COMMON RESOURCES

Whilst it is possible to show which of the specific resources were observed to be the most popular amongst the participants, the observed use of all the remaining specific resources can only be described as a *long tail*\(^{185}\).

The second aspect of the analysis of the use of specific resources as observed in the translation task of the Main Study is the *average* number of specific resources accessed during the task. When calculating the average number of specific resources per participant, the fact that certain resources were common to more than one participant was taken into account. The numbers of specific resources accessed by each participant were added up (totalling 163, see Figure 5.27) and divided by the number of participants (16), giving an average of 10.2 specific resources per participant.

\(^{185}\)The term *long tail* is used in statistics to describe the portion of the distribution having a large number of occurrences far from the central part of the distribution.
This figure of 10.2 specific resources per translator per task is almost exactly the same as in the 2009 study by Désilets et al. (2009:np), in which the average number of resources in translators’ tool boxes was found to be 10, thus supporting Désilets et al.’s statement that translators “use a wide variety of tools and resources” (Désilets et al. 2009:np.). The list of resources produced from their observation of eight subjects for fifty minutes is claimed to be “richer than any that has been previously reported” and “somewhat astounding” (ibid.), although the exact number of specific resources used by their subjects is not explicitly stated. This statement could, however, be validated by the extensive list of specific resources obtained from this study, which can be found in Appendix 7.

Although the number of resources in the translator’s tool box appears virtually the same as in Désilets et al.’s study (2009:np), this has to be treated with a degree of caution as the two studies differ up to a point in terms of what resources are included in the calculations. Whereas in Désilets et al.’s study any resource used during the study was taken into account, in the present study, certain translation aids were omitted from the calculation as they were treated as tools rather than resources or tools acting as resources (see section 4.1. for the distinction between tools and resources). These included spellcheckers, TM or MT integrated into a CAT Tool, which were applied to the task as a whole, rather than to solving particular translation problems. For example, if Google Translate was used for ad hoc terminology research, it was classified as a resource, whereas any machine translation engine, including Google Translate, when applied to the whole document for post-editing was treated as a tool186. In Désilets et al.’s study two spell checkers and (at least) three translation memories were included in the calculations; however it is difficult to establish whether this discrepancy is of great importance since Désilets et al. do not provide the exact number of specific resources used in their study187.

---

186 In total, twelve instances of the use of tools were omitted from the calculations, which would slightly alter the average number of resources in the translator’s tool box, from 10.2 to 10.9. However, this difference was not deemed to be consequential for the purpose of the generalisation that on average translators may access as many as 10 specific resources for a given translation task.

187 Désilets et al. report that their participants carried out a total of 75 resource consultations, but it is not clear whether these consultations include the common resources, i.e. those used by more than two participants. They list a total of 46 resources, but in some cases references are made to multiple resources being used without providing the exact number (e.g. private lexicons or bilingual documents related to the source text, where only examples are being given).
Nevertheless, to say that on average translators have ten specific resources in their tool box does not mean that there is little actual variation between the individual translators. On the contrary, further analysis shows that they did, indeed, vary significantly with regard to the number of specific resources they felt it necessary to access in order to complete the same translation task; it was observed that translators used between 0 and 28 specific resources, with standard deviation from the mean being 6.2. This range is far wider than the one given by Nord (2009:211), but it has to be taken into account that Nord’s study was carried out in the pre-Internet era. Nevertheless, the variations between translators in the number of resources they use are apparent. While some translators managed with a dictionary, search engine and a couple of parallel texts, others needed many more reference sources to meet their research needs. This extensive range with regard to the use of specific resources represented by the sixteen participants will be reflected in the resource volume category, which will form a vital part of the Translator Research Style Typology discussed in Chapter 7.

5.2.5. Frequency of resource use

Whilst it is important to discuss the resource types and the specific resources, it is also vital to look at the frequency of their use, including repeated use, to see the level of their exploitation by translators. In terms of online traffic, some resources experience heavy congestion, while others are visited less frequently. Similar observations can be made with regard to the frequency of resource consultations in this study, for both specific resources and resource types. Observations relating to frequency can also be made in relation to individual users, whereby some translators have few resources in their tool box but use them frequently while others use many resources, but less repetitively. These aspects will be discussed in this section by examining the frequency of the use of specific resources, followed by the frequency of use of the types of resources.

Specific resources - frequency of use

The frequency of use of specific resources is measured firstly by looking at the whole translation task and secondly, by calculating the frequency per research unit (RU), i.e. the actual lexical item in the source text being researched. Calculated for the entire task, the frequency of specific resource use takes into account all the instances of use of specific resources in all research episodes (as defined in section 3.2.3.) of all the participants. This cumulative calculation focusses on which specific resources were the most frequently used in the translation task of the Main Study.

The sixteen professional translators were observed to have accessed various specific resources a total of 865 times in the translation task of the Main Study. Figure 5.28 below shows the specific resources most frequently used by at least two translators, i.e. those considered to be ‘common’. Unsurprisingly, Google is ranked first with 282 instances of use, followed by Wikipedia\(^\text{188}\), ProZ Glossaries, Linguee and Wordreference. Interestingly, Google Translate is ranked sixth with 29 instances of use, which clearly indicates the adoption of MT by professional translators (see section 5.2.6.). Another observation worth mentioning is the popularity of the ‘social’ types of resources represented by two discussion fora - WordReference discussion forum and ProZ discussion forum which were consulted 11 and 10 times

\(^{188}\)Google followed by Wikipedia were also observed to be the most frequently used resources by Volanen (2015:46).
respectively. A breakdown of numbers in Figure 5.28 for each resource, showing how many translators contributed to the total numbers can be found in Appendix 8.

The most frequently used specific common resources

![The most frequently used specific common resources](image)

Figure 5.28 THE MOST FREQUENTLY USED SPECIFIC COMMON RESOURCES IN THE MAIN STUDY

The frequency of use of specific resources in the entire translation task also illustrates that the range of the number of steps taken by the participants to satisfy their research needs is very wide. As shown in Figure 5.29 below, translators took between 0 and 138 steps to address their specific research needs; that is, they consulted various specific resources between 0 and 138 times during the translation task. The average number of steps taken to satisfy all those needs is 54 consultations per participant. If we recollect from section 5.2.4. that the translators had on average 10.2 resources in their tool box, we can come to the conclusion that on average, each translator made 54 consultations using 10.2 resources during the translation task of a 412-word document.

![Number of steps, i.e. the frequency of consultations during the translation task in the Main Study](image)

Figure 5.29 FREQUENCY OF SPECIFIC RESOURCES USE PER TASK BY INDIVIDUAL TRANSLATORS IN THE MAIN STUDY

When the frequency of use of specific resources is measured per research unit, two calculations are of relevance: the number of steps per RU, i.e. how many times specific resources were used, and the number of specific resources per RU, i.e. how many different resources were used, to satisfy a particular research need. The difference between the number of steps and the number of specific resources used in a research
episode is illustrated below in Table 5.9. Whilst T1 took five steps to address the research need related to the research unit ‘WIR Bank in Switzerland’, only two resources were used during this research episode.

<table>
<thead>
<tr>
<th>Translator</th>
<th>T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research unit (RU)</td>
<td>WIR Bank in Switzerland</td>
</tr>
<tr>
<td>Research episode path</td>
<td>(1) Google → (2) Wikipedia → (3) Wikipedia → (4) Google → (5) Google</td>
</tr>
<tr>
<td>Number of steps</td>
<td>5</td>
</tr>
<tr>
<td>Number of specific resources</td>
<td>2 (Google and Wikipedia)</td>
</tr>
</tbody>
</table>

Table 5.9 EXAMPLE SHOWING THE DIFFERENCE BETWEEN THE NUMBER OF STEPS AND THE NUMBER OF SPECIFIC RESOURCES USED IN A GIVEN RESEARCH EPISODE

The individual differences regarding the number of steps per RU can be seen as an indicator of what I will call research intensity while the number of specific resources per RU contributes to the understanding of research diversity, both discussed in Chapter 7. Here, we are concerned with the cumulative numbers for each of the participants within the sample across all of their research units in order to see the overall tendencies. These numbers give the average number of steps and the average number of specific resources needed per research unit to solve a problem related to this unit, as observed in the translation task of the Main Study.

Figure 5.30 below shows both calculations, i.e. the average number of steps taken and the average number of specific resources consulted by respective translators per one research unit (RU). It can be seen that whilst at one end of the spectrum, T5 takes on average 1.5 steps to solve one translation problem, at the other end of the spectrum, T11 needs 4.5 steps to do the same thing. With respect to the number of specific resources per RU, the range is narrower as it includes repetitive use of specific resources.

Since this section concentrates on the common tendencies, the average number of steps per RU and the average number of specific resources per RU were calculated for all participants. On average, translators took 2.6 steps to resolve a translation problem and to do this, they used an average of 1.8 specific resources. Subjects in the study by Désilets et al. on average carried out 1.05 consultations per problem (Désilets et al., 2009:np), although it is not clear whether they calculated the total number of steps or the number of specific resources accessed per problem. Nevertheless, both figures obtained in this study (i.e. 2.6 steps per RU and 1.8 specific resources per RU) are higher than those reported by Désilets et al. (2009).
The types of resources - frequency of use

As with the frequency of use of specific resources, the frequency of use of types of resources is measured firstly by looking at the whole translation task and secondly, by calculating it per research unit (RU). When considering the frequency of use of the types of resources in the whole translation task, it becomes clear that although some types of resources are not abundantly represented in terms of the number of specific resources within the type, the type of resource itself is frequently used. For example, only two types of concordancers were used: Linguee and TAUS, but they were consulted 52 times by various participants during the task. Similarly, only two termbases were accessed in total, but they were consulted 14 times. Some types of resources were solely dominated by one specific resource. For example, the search engines category was solely represented by Google, despite the fact that many other multilingual search engines are available (see section 4.3.1.). Therefore, in the case of search engines, the frequency of specific resources use (see Figure 5.31) was the same as the frequency of the type of resource use, that is, 282. Similarly, knowledge-based resources such as encyclopaedias were dominated by the use of Wikipedia, which was consulted a total of 109 times out of 114 for all knowledge-based resources.

Another observation worthy of mention is the comparison of the use of monolingual vs. bilingual dictionaries. Scholars have argued that the use of monolingual resources is associated with the more experienced translators (see for example Jääskeläinen, 1989:186-188) while the less experienced translators tend to resort to bilingual sources. However, data obtained in the Main Study clearly suggests that the use of bilingual dictionaries is considerably more frequent even amongst experienced professionals, with 152 instances of bilingual dictionary use compared to 28 instances of monolingual dictionary use. These findings contradict the self-declared figures given by the same translators reported in section 5.2.3.

---

189 Linguee is a resource that is difficult to classify as it has two components: a specialist search engine and an editorial dictionary, but also behaves like a concordancer. In his latest Tool Box Journal, Zetzsche echoes this opinion saying “I’ve always had a hard time describing exactly what it is. Corpus tool? Search engine? Dictionary?” (Zetzsche, 2015:np). However, in his Tool Box Journal Zetzsche comes to the conclusion that “finally, in its latest incarnation, it’s clear just what it wants to be: a dictionary” (ibid). Due to the advanced stage of this study, however, Linguee will continue to be treated as a concordancer.

190 A related analysis of the directionality of research can be found in section 6.1.2., where source text vs. target text strategies as used by the participants in this study are discussed.
When measured per research unit, little difference was found between the frequency of the use of *types* of resources and frequency of *specific* resources. While the average number of specific resources accessed per RU, as noted earlier, was 1.8, the average number of types of resources per RU was 1.6 (see Figure 5.32). This would suggest that when choosing the next resource for consultation translators rarely choose the same type of resource and mostly switch to try another type of resource, with the exception of websites, which accounted for most of the differences between these two calculations.
As mentioned earlier in this section, while some translators may use many resources, they may not necessarily use them frequently. And vice versa - some might use few resources but use them very often. This is the case not only as far as specific resources are concerned, but also with regard to the types of resources. Below a breakdown of the frequency of resources use by type is presented for each type of resource.

Dictionaries, especially bilingual ones, were one of the most frequently used types of resources in the translation task as shown in section 5.2.3. However, translators differed significantly as to how frequently they used this type of resource to look for solutions to their translation problems and these individual differences will be explored further in Chapter 6. It suffices to say here that monolingual dictionaries were used by nine participants and the most intense use of monolingual dictionaries was T8 (see Figure 5.33). Bilingual dictionaries were used by thirteen participants and they were used more intensely than monolingual dictionaries, with T3 displaying the heaviest pattern of bilingual dictionary use.

Figure 5.32 AVERAGE NUMBER OF TYPES OF RESOURCES PER RESEARCH UNIT

Figure 5.33 THE USE OF MONOLINGUAL AND BILINGUAL DICTIONARIES BY INDIVIDUAL TRANSLATORS IN THE MAIN STUDY
Other types of dictionaries such as specialised dictionaries, dictionaries of idioms or synonyms were only used sporadically and, therefore, cannot be discussed here in terms of the frequency of their use.

With regard to other types of common resources, i.e. those shared by at least two participants, and the frequency of their use by the participants in the Main Study, Table 5.10 below shows which types of common resources were used most frequently, how many participants consulted them and who were the most frequent users of these particular resources.

<table>
<thead>
<tr>
<th>Type of common resource</th>
<th>Number of participants using common resources</th>
<th>Most frequent users of common resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search engines (Google)</td>
<td>15</td>
<td>T9 (56 consultations) T11 (50 consultations)</td>
</tr>
<tr>
<td>Knowledge-based resources</td>
<td>14</td>
<td>T3 (26 consultations) T11 (25 consultations)</td>
</tr>
<tr>
<td>Web pages</td>
<td>11</td>
<td>T11 (50 consultations)</td>
</tr>
<tr>
<td>Concordancers</td>
<td>7</td>
<td>T1 (20 consultations) T10 (17 consultations)</td>
</tr>
<tr>
<td>Discussion fora</td>
<td>6</td>
<td>T15 (7 consultations)</td>
</tr>
<tr>
<td>Glossaries</td>
<td>5</td>
<td>T1 (17 consultations) T3 (15 consultations)</td>
</tr>
<tr>
<td>Term banks</td>
<td>4</td>
<td>T7 (6 consultations)</td>
</tr>
<tr>
<td>Online documents</td>
<td>4</td>
<td>T11 (2 consultations)</td>
</tr>
<tr>
<td>Non-customised MT (Google Translate)</td>
<td>3</td>
<td>T11 (20 consultations)</td>
</tr>
<tr>
<td>Thesauri</td>
<td>2</td>
<td>T1 (2 consultations)</td>
</tr>
</tbody>
</table>

Table 5.10 MOST FREQUENTLY USED COMMON RESOURCES OBSERVED IN THE MAIN STUDY

The frequency of use of other common resources such as glossaries, termbases, concordancers, discussion fora, non-customised MT, search engines, knowledge-based resources and web pages can be found in Appendix 9. As an example, although fifteen out of sixteen participants used Google search engine, some rely on Google more than others. While T9 used it 56 times, T5 and T14 only accessed it once. Similarly, knowledge-based resources such as Wikipedia, although used by fourteen out of sixteen participants, were accessed 26 times by T3 and once by T5.

5.2.6. The patterns of machine translation (MT) use

Taking into account the importance of developments in the area of machine translation mentioned in Chapter 4 but without engaging in the wider debate around the subject, MT will only be discussed here in the context of the aims of the present study. The discussion will focus on the use of this
technology by professional translators as self-declared and as observed, illustrating how it is used by different user groups and individuals. Data concerning the use of machine translation was elicited from both study components, the Global Survey (self-declared) and the Main Study (self-declared and observed).

In his recent Tool Box Journal, Zetzsche (2014:np) provides some statistics on the use of MT by professional translators. The data was obtained from MemSource, a cloud-based translator’s workbench which makes MT available from within its suite of tools and can track how many translators actually use it. These statistics, based on 30,000 user accounts (Zetzsche, 2014b), revealed a surprisingly high frequency of MT use and prompted much debate in the translation communities, e.g. see LinkedIn (2014:online).

According to MemSource, MT is used by 46.2% of subscribers, with just over 98% of the 46.2% using public, non-customised engines such as Google Translate or Microsoft Bing Translator, which store and reuse the submitted segments. Only less than 2% of those using MT through MemSource chose to use non-public, customised and secure engines such as Apertium191, Asia Online, KantanMT192, LetsMT193, MoraviaMT194, PangeaMT195, Systran196, and Tauyou (Zetzsche, 2014b) (see Figure 5.34).

![Figure 5.34 THE USE OF MACHINE TRANSLATION IN MEMSOURCE](image)

Fontes (2013) reports even higher numbers in the European Commission’s MT survey. A year after MT was rolled out across all language departments of the Directorate-General for Translation, 535 of the 763 translators who responded to the European Commission’s MT survey (approximately 70%) reportedly used MT in their translation work and 16% said they tried MT, but did not use it. The remaining translators said they had never tried MT in their translation work at the European Commission.

In the present study, the number of translators using machine translation in their work, as mentioned above, was investigated in both study components and two types of data - self-declared and observed -
were collected. The self-declared data collected as part of the questionnaires from both study components (the Global Survey and the Main Study) will be considered first.

In the Global Survey\textsuperscript{197} 53\% of the respondents reported using non-customised MT whilst 23\% said they use customised MT solutions. In the Main Study, three quarters of the much smaller sample of sixteen translators declared using non-customised MT and 37\% declared using customised solutions (see Figure 5.35).

![Figure 5.35 SELF-DECLARED USE OF MACHINE TRANSLATION COMPARISON OF GLOBAL SURVEY AND MAIN STUDY DATA](image)

In both questionnaires, a 5-point Likert scale was used to determine to what degree translators use MT. For this calculation, the answers from the four points on the scale, i.e. \textit{mostly}, \textit{frequently}, \textit{sometimes} and \textit{rarely} were aggregated into the \textit{yes} category and the answers from the last point on the scale, i.e. \textit{never} are represented in the \textit{no} category. Figure 5.36 below shows the distribution of the answers on the 5-point Likert scale in both the Global Survey and the Main Study.

![Figure 5.36 SELF-DECLARED USE OF CUSTOMISED VS. NON-CUSTOMISED MACHINE TRANSLATION IN GLOBAL SURVEY AND MAIN STUDY](image)

\textsuperscript{197}This question was answered by 527 respondents (96\% of the total sample of 540).
To summarise the above findings, Zetzsche’s statement below seems fitting:

> There has been a chorus of voices in the last few years, especially in the MT community, claiming that even though so many translators are complaining about MT, a relatively large percentage of them are “secretly” using it. With almost half of all Memsource users proven to be using machine translation, I think we can say that there is some credence to that notion.

(Zetzsche, 2014:np)

In the translation task of the Main Study, participants were asked to use resources they would normally use during their work as far as this was possible. Out of the sixteen participants, four used MT during the translation task. One had reported using MT (Google Translate) but did not actually use it during the task.

Out of the four participants, three used MT for ad-hoc terminology or phraseology research, corresponding to the search and discovery use of MT identified by Van Der Meer & Ruopp (2014) and two used MT to translate the whole text from within their CAT Tool, corresponding to the post-editing function of MT (ibid.). As mentioned in section 4.1. and elsewhere, it was decided that reference aids applied to the whole task should be treated as tools while those used to solve a particular translation problem should be treated as resources. This was also the case with machine translation and, therefore, the use of MT as a post-editing tool is not being investigated here. Based on this, three ‘user groups’ were identified: the Search and Discovery users, the Post Editors and the Non-users which will be described below.

**The Search and Discovery users** (T5, T11 and T14) all used Google Translate (GT), although they differed in the purpose and the extent of its use. T5 used GT as the main resource in the whole translation task, consulting it on eight occasions (73% of all consultations) by checking how particular words or phrases were translated by the GT engine. In the post-task questionnaire, T5 said that “Google Translate is a very useful tool when you master the language, that is, when you are able to say that its translation was no good. In fact, sometimes I go to it just out of laziness; not having to think too much” (see Appendix 1 for the full post-task questionnaire). T11 was the most prolific user of GT, interrogating the system for single words, phrases and whole sentences on twenty occasions (see Figure 5.37 for examples). However, GT was not T11’s main resource as GT consultations accounted for 15% of all her consultations. T11 was enthusiastic about Google Translate saying it was “well fed” for the legal texts which she mainly translates, although she did make a reference to it as a being a “cheating tool”. When asked in what way she finds Google Translate useful, she responded that it helps her to find the right collocations and that she uses it solely for repacking procedures (i.e. reformulation) rather than comprehension.
T14 used GT as a resource on only one occasion, but it is interesting to see that he also used it in the post-editing mode. Also noteworthy is the fact that, whilst using GT, T14 made a number of corrections within the copied segment which were subsequently submitted and incorporated into the learning mechanism of GT.

Two translators (T8 and T14) were identified as Post Editors. They both used machine translation as a translation tool applied to the task as a whole\textsuperscript{198}. T8 used the freely available translation memory software Wordfast Anywhere, which was set to use Microsoft’s machine translation database. He admits that, for him, post-editing MT output is quicker than translating from scratch, which is interesting when taking into account that the target language in this case was Indonesian, which is not one of the major MT languages for which there are petabytes of data on which to train MT engines. T14 used SDL BeGlobal Community from SDL Trados although he also reported setting MT to Google Translate in SDL Trados, depending on language combinations.

The third, Non-Users group, accounted for 75% of the sample. T1 reported using GT on occasions during her normal work but did not actually access it during the task.

Based on the high reported use of MT in the present study and the fact that 25% of the sample of the Main Study used MT in one way or another in the translation task, the question is no longer whether professional translators use MT, but how they use it, i.e. whether it is used as an ad-hoc research tool or whether it is applied to the whole document with subsequent post-editing. As demonstrated in the present research, two of the five uses of MT put forward by Van Der Meer & Ruopp (2014), namely search and discovery and

---

\textsuperscript{198} Although this study focusses on the use of resources and MT used in a post-editing mode is classified as a tool, it is nevertheless described here due to the lack of empirical MT user case studies at the point of writing. Therefore, it is considered that this information can be of interest for future researchers on MT use amongst professional translators.
post-editing have been employed by four out of sixteen professional translators in the present sample. Zetzsche (2015b) also confirms that “many of us are looking for a deeper and different integration of machine translation into our translation environments”. While some translators are interested in using machine translation through the traditional route of PEMT (post-editing of machine translation), overall there is more interest in having machine translation be just one of the resources in our translation environments. However, more research is needed in this area to establish how professional translators are adopting MT technology and in what way they are utilising this technology.

5.3. Research distribution - The How

Having discussed who took part in this study and examined the patterns of participants’ resource use (the what) during their research activities, we will now discuss how these activities were distributed within the translation task of the Main Study. The following sections will be dedicated to exploring the progression of the sixteen professional translators through the translation task as they solve various translation problems. This exploration will focus on the micro patterns of research distribution (patterns within a research session, i.e. all research activity related to one research unit) and macro patterns of research distribution (within the translation task). In this context, as mentioned in 1.3., translation task is understood to consist of a succession of translation episodes (interactions with the source and target texts) and research episodes (interactions with online resources).

5.3.1. Micro-patterns of research distribution

The micro-patterns of research distribution refer to research activities observed within individual research episodes. The analysis of these micro-patterns reveals the complexity of translation-oriented research activities within those episodes and shows how momentary the components of these episodes can be. In order to understand these research activities in the context of research sessions, we need to firstly bear in mind the distinction between the interactions with texts, which are described in terms of translation episodes, and the interactions with external resources, referred to in terms of research sessions, research episodes and research steps (see section 3.2.3. for definitions). Secondly, these concepts will be briefly recapped by way of example in order to facilitate the comprehension of the data presented. The diagram below represents an example of how a research unit (Bitcoin) was researched during the translation task by T3 (see Figure 5.38).

---

199Based on input from professional translators to Zetzsche’s call for ideas about new directions in translation technology.
In this particular research session, the research unit is ‘Bitcoin’ and the research session that was carried out to satisfy the research needs related to this research unit consisted of three research episodes. All three research episodes happened at different times in the translation task but because they concerned the same research unit, they form a research session. Episode one and three comprised three research steps, whilst episode two had only one research step. It must be noted, however, that most research sessions consisted of one research episode only. Research sessions consisting of more than one research episode were observed only on nine occasions, which is 3% of all research sessions. It is also important to bear in mind that the concept of research session is tied to the concept of research unit (RU), which refers to the word or phrase being researched (see 3.2.3. for a definition). So, the research session reflects all the research steps in all research episodes being taken to address research needs associated with a particular research unit. Below, four micro patterns will be examined by looking at the smallest units of analysis, the research steps, their relationship to research units, followed by considering the larger units, translation and research episodes. These analyses relate to the common tendencies observed in the Main Study sample and a more in-depth analysis of individual differences pertaining to these observations will be carried out in Chapters 6 and 7.

Firstly, the length of research steps was calculated for each of the individual translators in order to provide important information about the pace of research progression. The steps from which research episodes were made were found to be very short, which means that individual resources were consulted very briefly. The average duration of a research step for all participants, excluding T2, lasted 20”, which means that on average, it takes 20” to consult one specific resource. However, translators varied significantly in relation to the length of these steps. Some took on average as little as 12” to consult a resource (T9) while others took as long as 38” (see Figure 5.39). The longest observed research step in all observed screen recordings was 1’47” and the shortest was just 2”. Table 5.11 below summarises these calculations.

<table>
<thead>
<tr>
<th>Description</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average duration of research step for all participants</td>
<td>00:00:20</td>
</tr>
<tr>
<td>The range of average duration of research steps</td>
<td>00:00:12 – 00:00:38</td>
</tr>
<tr>
<td>The longest observed research step</td>
<td>00:01:47</td>
</tr>
<tr>
<td>The shortest observed research step</td>
<td>00:00:02</td>
</tr>
</tbody>
</table>

Table 5.11 THE LENGTH OF RESEARCH STEP
Secondly, the number of research steps per research unit was considered in order to determine how intensely a research unit had been researched in all research episodes. This was discussed in the context of the frequency of specific resources use (see section 5.2.5.). It was observed that while some translators did not engage in the lengthy pursuit of an acceptable answer and averaged 1.5 steps per research unit (T5 being the least intense), others needed on average 4.5 steps to achieve the same goal (T11 being the most intense) (see Figure 5.30).

The highest number of observed steps in a research session composed of four research episodes was observed for T11, who took 22 research steps in all four research episodes to research the term ‘digital currency’. The whole research session lasted 4’17” and its path looked as follows:

Research Session No 1 (T11):

Research Episode 1: Google → Wikipedia → web page4 → web page5 → web page2a → web page2b → Google
Research Episode 3: Google → web page4 → Google → web page2 → web page2b (abandoned) → Google → web page4 → web page7
Research Episode 4: Wikipedia → Google → web page2

The longest single research episode was observed for T1, who researched the word ‘falter’ for 4’17”, taking 11 steps. This was followed by T15, who researched the term ‘disruptive’ for 4’16” taking 13 steps and another term, ‘haircut’ in 4’ taking 15 steps.

Although some episodes were exceptionally long and complex, it is also worth pointing out that 34% of all research episodes consisted of just one-step. Translators varied considerably with regard to the proportion

---

200 The individual websites consulted were numbered for easier identification during the analysis.
of one-step consultations, ranging from 0% in the case of T12 to 70% in the case of T13 (see Figure 5.40 below). The one-step consultations were predominantly associated with consultations in bilingual dictionaries although one-step Google consultations were also common. This will become important when calculating research intensity is section 7.1.1.3.

The one-step consultations were predominantly associated with consultations in bilingual dictionaries although one-step Google consultations were also common. This will become important when calculating research intensity is section 7.1.1.3.

The third micro-pattern of the distribution of research activity refers to the switching between research episodes and translation episodes in which translators temporarily suspend the interaction with the translated text in order to engage in a completely different task of interaction with external resources, i.e. information seeking. The switching from translation episode to research episode can be repeated numerous times during a translation task, sometimes surprisingly frequently. This, at times, can make an otherwise natural process appear as a rather disjointed activity. The ramifications of such switching from a cognitive point of view must be wide-ranging and far-reaching; however, because of their cognitive nature, these are not a subject of investigation in this study. In the present work, this switching between translation episodes and research episodes has been captured quantitatively and provides insight into the distribution of the translation episodes and research episodes within the translation task. To the best of my knowledge, this particular area of switching between what I have called research episodes and translation episodes has only partially been addressed before in a quantitative manner (see e.g. Nord, 2009; Enríquez Raído, 2011) and offers potential for future larger-scale research.

The average length of translation episodes is 1’ 27”, excluding T2, who had only one translation episode which lasted 30’30”. Figure 5.41 below shows average lengths of translation episodes for each participant. The shortest average translation episode was just 41” (T3) and the longest was 2’39” (T6). The longest translation episode observed was 5’32” (T8) and the shortest was just 3” and was observed on several occasions.
With regard to the average length of research episodes, it was observed that they varied, from 35” (T3) to 1’48” (T1), with the average length of a research episode of 1’1” for all participants excluding T2 (see Figure 5.42 below). As mentioned previously in relation to the number of steps per research episode, the longest single research episode was observed for T1, who researched the word ‘falter’ for 4’17″, followed by T15 who researched the term ‘disruptive’ for 4’16″ and another term, ‘haircut’, for 4’. The shortest research episodes were one step consultations, mostly in bilingual dictionaries, lasting for 2” (e.g. T3).

To summarise the discussion pertaining to the length of translation and research episodes, with the average of 1’27″ for translation episodes and 1’01” for research episodes, it is clear how frequently switches between translation and research episodes can occur. The average figures for research episodes obtained in the present study are in line with the figure of 59” obtained by Enríquez Raído (2014:156) for what she calls ‘session lengths’ in her first embedded task. Volanen (2015:50) comes close with the average
of 1’26” observed in three participants’ IS contexts (called research steps in the present study). Enríquez Raído’s average of 1’51” obtained in a sample of translation students in her second embedded task (Enríquez Raído, 2014:256) would suggest that the task difficulty might influence the length of research steps, although it is difficult to compare the text used in the above mentioned study. Comparing these figures to Nord’s pre-Internet study, in which translation process was observed to be interrupted on average every 3.5 minutes (Nord, 2009:214), would imply that the translation episodes are getting shorter and the consultations more frequent with the use of online resources (as opposed to paper ones).

However, the degree of switching can differ considerably between translators, something that is masked by averages. Some translators’ work is characterised by many short translation episodes while others exhibit patterns of longer periods of engaging with the texts. For example, on the one hand, T3 displays the shortest length of average translation episode (41”) and the highest number of such episodes (65). On the other hand, T6 exhibits the longest length of average translation episode (2’39”) and one of the lowest numbers of such episodes (12) (see Figure 5.41 and Figure 5.43). Therefore, the switching pattern depends not only on the length of the episodes but also on the number of these episodes during the translation task.

![Figure 5.43 OBSERVED NUMBER OF TRANSLATION EPISODES](image)

### 5.3.2. Macro-patterns of research distribution

As mentioned in the previous section, capturing the interactions with external resources within the translation process proves to be a challenging task. Apart from the micro patterns of research distribution within a research episode, we also need to consider how these research episodes are distributed within a translation task, thus forming macro patterns of research distribution. The macro perspective, therefore, adds another dimension to examining the distribution of research activities.

From the macro perspective, research can happen at any time during the translation process. Although the screen recordings show that the majority of research is carried out immediately after the identification of a particular problem and is distributed in short steps within research episodes during the translation task, some translators’ research activities can be distributed less evenly, forming larger clusters at certain points
in the process. Therefore, research activities within this translation process are not always clustered around the research unit; translators perform other research actions that are difficult to capture, such as what I have called backtracking, indirect research, prospective research and retrospective research. Below, these five macro patterns of research distribution are discussed in a qualitative fashion according to the observations made in the screen recordings which form part of the Main Study.

**Clustering**

It was noted during the analysis of the screen recordings that some translators cluster their research at different points in the translation task. Four main categories of this clustering, which I have called front loaded, mid-task, end-loaded and continuous research have emerged and these will be discussed in the context of Carl et al.’s taxonomy of human translation styles (Carl et al., 2011:online) in which the following phases in translation process are distinguished (see also Englund Dimitrova, 2005:21-22):

- *initial orientation*, which functions as an initial text-planning phase
- *translation drafting*, in which the translation is drafted
- *revision*, in which the text is reviewed and translation choices refined

**Front-loaded research** is characterised by the fact that a large bulk of the research happens during the initial orientation phase, i.e. at the beginning of the translation task, before starting to produce a translation. In this type of clustering, translators peruse parallel texts to read around the subject and get an idea of what terminology might be useful. Often they look up key terminology that ‘jumps out’ from the text. Translators who perform systematic orientation also tend to put more emphasis on front-loaded research. T9, T10, T11 and T15 were observed to conduct such research.

**Mid-task research** is characterised by a large cluster of research being carried out in the middle of the task, in the drafting phase. T6 admitted to routinely conducting mid-task research as “very often the text itself brings more insight into the subject matter [and] the explanation of the terms and concepts that are unfamiliar come further in the text”. Therefore, she found her research to be more efficient once she has “grappled with the text”. Two translators, T6 and T13, conducted mid-task research.

**End-loaded research** is characterised by a large cluster of research being carried out at the end of the task, in the revision phase. T16 is a good example of a fast drafter who carries out a heavy revision during which terminology is further researched and corrected. Head starters, i.e. those who skip the initial orientation phase, are often found to be heavy revisers. Two translators, T14 and T16, carried out end-loaded research in the present study.

**Continuous research** is characterised by an even spread of research, from start to end. Seven translators, T1, T3, T4, T5, T7, T8 and T12, conducted such research during the translation task. Although four types of research distribution were identified in this study, the main modus operandi is based on continuous research. However, some translators evidently put more weight at the beginning, middle or back end of the translation task. This adds to the complexity of the structure of translation-oriented research. These clustering patterns can be linked to translation styles found by Dragsted & Carl (see section 2.6.), especially with regard to the drafting phase and the revision phase (Dragsted & Carl, 2013). It could be said

---

201 References to Carl et al.’s (2011) taxonomy made in this study will refer to different phases of the translation process in which research occurs.

202 Quoted from the post-task questionnaire which can be found in Appendix 1.
that the front loaders correspond to Dragsted & Carl’s translators with an extended orientation phase, the end-loaders could be Dragsted & Carl’s end-revisers and the continuous-researchers in the present study could be Dragsted & Carl’s online revisers. However, these correspondences can only be treated tentatively as Dragsted & Carl did not allow the use of external resources in any form during the experiment. Therefore, any activities observed during the orientation, drafting and revision phases would have been carried out solely using internal resources.

**Backtracking**

Backtracking in the context of this research refers to any activity that is characterised by flicking backwards and forwards between the text and resources. It can refer to going back to the previous resource after a translation episode, usually to double-check, or it can be switching back to the text, usually to double check the context in which a word occurs. While backtracking can be discussed in terms of switching between texts and resources within a research episode (see section 5.3.1.), here, the emphasis is on the fact that the switching is a backward movement and that it occurs repeatedly during the translation task. The most prominent ‘backtrackers’ were T3, who was observed to backtrack ten times during the task, T15 who backtracked six times and T10 who did it five times.

Establishing the importance of backtracking in translation process research can be problematic as it is usually a momentary action, sometimes lasting seconds; however, it indicates an important aspect of translators’ interaction with online resources. It would appear that backtracking, whether to the source text or to the resource, fulfils the function of an ‘aide-memoire’ to which translators can have quick and easy access.

Interestingly, when asked to rate their retentive memory in the post-task questionnaire, two of the backtracking translators rated themselves as 3 and one rated themselves as 4 out of 5, 5 being the highest. This was not lower than the average reported for all the participants, which was 3, thus indicating that the self-declared level of retentive memory did not correspond with the level of observed backtracking. Therefore, we can ask whether backtracking is a phenomenon that characterises the workings of those brains with a lower capacity for retention or whether it is an evolutionary response to technology, whereby easy access to information somehow permits us to treat the virtual information warehouse as a memory aid. Clearly more research is needed in this area, but this can be left in the hands of psycholinguists.

**Indirect research**

Indirect research can be defined as finding potential terms while reading parallel texts, not on impulse after encountering a particular research need. As mentioned in section 2.6., studies in information behaviour have identified similar behaviour as information encountering or incidental knowledge acquisition. In such instances, information is serendipitously encountered, recognised as potentially useful and stored for later retrieval in the translator’s memory or saved immediately in ‘extended’ memory such as a termbase or a glossary.

A prime example of this kind of research is given by T6, who was observed to utilise nine solutions she had found while reading background texts. She was observed to insert a solution into the target text while mentioning that she had seen it in a certain parallel text and therefore, was confident she could use it.
Indirect research is extremely hard to capture unless the participant clearly indicates in their accompanying narrative that a potential translation problem had existed but was solved by employing a solution found while reading parallel texts. If not verbalised, this type of research cannot be captured, even with the help of eye-tracking as it would simply appear as if a translator were using their internal resources. Indirect research demonstrates the importance of engaging with parallel texts during the research process, especially in unfamiliar domains.

**Prospective research**

Prospective research refers to the researching of a term long before the research unit appears in the text. It demonstrates that research can take place even before a potential research unit in the source text has been identified as requiring research. The motivations for this kind of research include an initial skimming or reading of the text, anticipating that a term will need researching and linking this future need with the present need on the basis of some connection between the research units. For example, while researching institutions such as ‘WIR Bank’ and ‘Ithaca Hours’, which occur in one paragraph, it could be considered beneficial to research the related ‘Liberty Reserve’s LR’, which occurs in the third paragraph. This is precisely what T11 has done. Prospective research is differentiated from the front-loaded research distribution by the individual nature of the query.

**Retrospective research**

Retrospective research is the reverse action of prospective research, i.e. a research unit is researched well after it was first encountered, after more than one translation episode has taken place, as a result of second thoughts about something in the ST that requires further research. Retrospective research is considered to be part of the research session for a particular research unit to which it refers, but it is distinguished as a separate episode within the research session.

For example, T6 researched ‘Bitcoin’ in two research episodes:
Research Episode 1: Google → Wikipedia (relating to world knowledge – i.e. what is Bitcoin)
Research Episode 2: Google → Google (Query modification) → web page4 (relating to word knowledge, i.e. establishing whether to capitalise or not).

During retrospective research, translators may be consulting the same resource as previously or they might use a new resource. A total of seven translators were observed to carry out retrospective research for the following research units:
- T1 - rise
- T3 - established, bitcoin
- T6 - bitcoin
- T11 - digital currency, rise
- T13 - currency
- T15 - bitcoin, mining process
- T16 - fiat

---

203 If only one translation episode took place after the original research episode, then this action is called backtracking.
To summarise, the distribution of research within the translation task is a complex issue, which, as far as I am aware, has not been studied before from the information behaviour point of view, although many early process-oriented studies (see Enríquez Raído 2011:17-20 for a detailed account) have reported on the ‘non-linearity’, the ‘iterativeness’ and ‘parallel processing’ in the translation process (e.g. see Séguinot, 2000:146). It would appear that there is little understanding of how exactly translators interact with online resources from the process perspective and how this process relates to their interactions with the source and target texts. As the present research demonstrates, the nature of research episodes is highly composite, there is frequent switching between the two types of interactions and the distribution of research episodes during the translation task is often irregular. As a result, to the external observer the research activities could be perceived as interrupting the translation activity (characterised as interactions with the texts), and yet translators manage to integrate these two activities in a continuous flow. Although the present research is not concerned with the potential disruptiveness of the research activities, certain behaviours have been observed which suggest that individual translators differ in the way they cope with the disruptive potential of the interactions with resources. While some translators start off by careful planning of the items to be researched in order to prevent them from wandering off their path, others ‘head start’ and subsequently meander through various resources spontaneously absorbing the interruptions into their translation process. Still others might consciously avoid excessive research and the subsequent disruption or simply manage it by fast and shallow processing. It has also been observed that some translators work more intensively than others, i.e. they manage to undertake more research activity in a given period of time than others, who work at a more leisurely pace. They might not only research more terms than others but also consult more resources than others in the same time. All these factors point to the fact that translators do not only have different translation styles (see Carl et al., 2011; Dragsted & Carl, 2013) but that they also develop their own research styles. The non-linearity attributed to the higher level of an individual translator’s level of competence or experience has been previously endorsed by some (e.g. Krings, 1986 in: Kiraly, 1995:48; Pöntinen & Romanov, 1989) and refuted by others (e.g. Kiraly, 1995:87; Jensen, 1999:111) leading Enríquez Raído (2011:20) to consider that “it may of course also be a question of different translating and/or cognitive styles”.

Over the course of their careers, translators develop their own way of finding the information they need, which can depend on many factors such as the training they have received, the availability of resources in their language pair or their level of interest in new technologies and new language resources. Furthermore, certain aspects of translators’ online information behaviour could be attributed to their innate character traits such as their personality (Heinström, 2005) or social style (Merrill & Reid, 1981). This combination of factors shapes the way translators approach their research for translation purposes, which manifests itself in a variety of behaviours. In the following two chapters, these behaviours will be examined in detail in an attempt to find common patterns and to group similar behaviours together. In doing so, the emphasis will be shifted from group patterns of behaviour to individual differences.
6. Individual differences – Resource Type User Typology (RTUT)

While Chapter 5 was mainly focussed on group profiles and common tendencies, observed with regard to the translation-oriented online research behaviour, Chapters 6 and 7 centre around individual differences between professional translators. As already noted, in the past, process-oriented research has mostly relied on ‘between-groups’ designs with the aim of comparing professionals with semi- or non-professionals. The present study concentrates solely on professionals and provides insight into their practices from a ‘within-group’ perspective. In the following two chapters, some individual differences will be outlined and clustered to present these differences in a schematic, but also, a systematic way.

In the course of analysing the data, it become apparent that the observations about translators’ information seeking behaviour encompass two aspects of translators’ online information behaviour. The first is concerned with resource behaviour, i.e. the type of resources that are dominant in each translator’s online interactions and the way in which these resources are accessed. The second relates to research behaviour, i.e. the way research is conducted in terms of quantities, timings and related qualitative factors such as diversity, depth of search or the shape of research paths, thus indicating research style. Therefore, two respective typologies are emerging from the data. One I have called the Resource Type User Typology (RTUT), and the other Typology of Translator Research Style (TTRS).

Despite the constraints and limitations of the research design of the current study, it is believed that the resource dominance and the research style of professional translators will manifest themselves through the collected data and that the taxonomies presented will duly capture and represent the information contained in this data. As Saldanha & O’Brien aptly observe with regard to “observations of individual differences within groups” in the translation process research, “no matter how controlled the experiment, the translator’s individual personality, experience, attitude and so forth will have an impact” (Saldanha & O’Brien, 2014:147, original emphasis). In the rest of the chapter, a Resource Type User Typology (RTUT) will be developed by first establishing the categories that feed into it, followed by the presentation of the typology. The RTUT is built upon the Main Study observations of the sixteen participants, some of which have already been highlighted in Chapter 5 while discussing common tendencies in the analysed sample. The Typology of Translator Research Style (TTRS) will be discussed in Chapter 7.

6.1. Categories feeding into the Resource Type User Typology

This section focusses on the individual differences between translators in the use of online resources which were demonstrated during the translation task of the Main Study. Unlike the observations in Chapter 5, which were restricted to discussing the use of the types and the specific resources, the observations in this chapter will encompass the examination of the use of online resources from other perspectives and will require a different classification of resources as mentioned in section 4.3.4. This is required in order to capture the individual differences in the use of resources in search of patterns, which will lead to establishing the profiles of resource behaviour. In what follows, the use of resources by the sixteen participants will be explored in relation to the three categories and their sub-categories (presented in Table 6.1 below) which will provide the building blocks for the Resource Type User Typology (RTUT) in section 6.2. These categories have been designed for the purpose of the present research, with the exception of ‘the nature of resources’ category, which is partly based on the distinction made by Désilets et al. (2009).
The three categories are intended to capture certain behaviours related to the use of online resources which are believed to be helpful in establishing the resource user types.

<table>
<thead>
<tr>
<th>RTUT CATEGORY</th>
<th>CATEGORY CHARACTERISTICS</th>
<th>MT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 THE NATURE OF RESOURCES</td>
<td>TERMINO-LEXICOGRAPHIC TEXT-BASED</td>
<td></td>
</tr>
<tr>
<td>2 RESEARCH DIRECTION</td>
<td>ST-ORIENTED</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT-ORIENTED</td>
<td></td>
</tr>
<tr>
<td>3 RESEARCH STRATEGY</td>
<td>TOP-DOWN</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BOTTOM-UP</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1 CATEGORIES AND SUB-CATEGORIES OF THE RESOURCE TYPE USER TYPOLOGY (RTUT)

The first RTUT category, the nature of resources, distinguishes between those translators who prefer using dictionaries, glossaries and term banks and those who predominantly use resources such as parallel texts (web pages) and other text-based resources. The third member of this category, MT, relating to the use of machine translation, is an add-on category and will only be discussed in this chapter where relevant.

The second RTUT category, the research direction, reflects the search orientedness of a translator in terms of whether they performed searches in the source or target language. This is tied to the use of particular resources as source-oriented searches were mostly performed in the termino-lexicographic resources and target-oriented searches were performed mostly in text-based resources.

The third RTUT category, research strategy, differentiates those translators who invariably initiate their research with a keyword in a search engine, from those who start from a resource they know, usually a termino-lexicographic one. All three RTUT categories will be discussed below in detail.

6.1.1. The nature of resources

The first step of the analysis leading to establishing resource type dominance involves looking at the nature of the resources used in the translation task by the sixteen translators. For the purpose of this part of the analysis, all the types of resources used by subjects in the translation task as represented in the a posteriori classification were divided into two groups, termino-lexicographic and text-based as elaborated on in section 4.3.3.

As noted, the boundaries between termino-lexicographic and text-based resources are not always easy to demarcate, especially with the current technological developments such as those in computational lexicography discussed in Chapter 4. However, for the purpose of this study, the resources used in the present study have been grouped into termino-lexicographic and text-based as presented in Table 6.2 below.

---

204 ‘Ask an expert’ was not included in the analysis of individual differences.
One type of resource, namely machine translation as used in a search and discovery mode (i.e. as a resource), did not fall into either of these categories. While Désilet’s distinction between the termino-lexicographic and corpus-based (text-based) resources is helpful and has been partly adopted for the purpose of the present research, in the context of the present work it is incomplete. Since machine translation engines\(^{205}\) are trained on data that can be found in both types of resources, text-based and termino-lexicographic, the output obtained from them is a statistical probability resulting from the analysis of both types of data. Therefore, MT was included as a third, add-on category in the classification of resources according to their nature. For the purpose of analysing the data, translators were first categorised as belonging to either termino-lexicographic, text-based or mixed category (indicating a mixed use of both termino-lexicographic and text-based), and then, if machine translation was used, they were also assigned MT as a secondary category.

Both termino-lexicographic and text-based resources were used by fifteen out of sixteen translators in the task and the distribution of their use was calculated in terms of the number of specific resources and the frequency of use of these specific resources based on tertiles. So, if the distribution, whether in relation to the number or frequency, is that 1/3 of the resources used are termino-lexicographic and 2/3 are text-based, this indicates that text-based resources are dominant. Conversely, if 2/3 of resources used are termino-lexicographic and 1/3 are text-based, a termino-lexicographic dominance is established. In the Mixed-type, there is no clear dominance of either termino-lexicographic or text-based resources. The blue colour indicates text-based resources while the red colour indicates the termino-lexicographic ones. Table 6.3 below shows how the resource dominance is established based on the percentage of use of resources

\(^{205}\) Only statistical machine translation was observed to be used in this study and therefore, the term ‘machine translation’ is used throughout to refer to statistical machine translation.
according to their nature. The two remaining RTUT categories (research direction and research strategy) will follow the same procedure for establishing dominance.

Table 6.3 THE PROCESS OF ESTABLISHING RESOURCE DOMINANCE

<table>
<thead>
<tr>
<th>type of resource</th>
<th>percentage of use</th>
<th>established dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0%-33%</td>
<td>34%-67%</td>
</tr>
</tbody>
</table>

Table 6.3 below illustrates how all individual participants were classified according to the distribution of termino-lexicographic and text-based resources used in the translation task. For example, T3 used a total of 4 termino-lexicographic resources (4/9, i.e. 45%) and 5 text-based resources (5/9, i.e. 55%). With regard to the frequency of use, T3 consulted the 4 termino-lexicographic resources 53 times during the task (53/99, i.e. 54% of all T3’s consultations during the translation task) and the 5 text-based resources 46 times (46/99, i.e. 46% of all T3’s consultations). Therefore, because T3 used a mixture of termino-lexicographic and text-based resources, falling within the 34%-67% range (see Table 6.3), T3 was assigned to the mixed category. As illustrated in Table 6.4, five translators T8, T9, T10, T13 and T14 were classified differently in each of the perspectives. This is because they might have used fewer resources of a particular nature, but used them more frequently. Both of these perspectives will be taken into account in the final classification leading to the Resource Type User Typology.
A summary of which individual translators were classified into the termino-lexicographic, text-based and the mixed sub-categories is presented below in Table 6.5. As in Table 6.4, the classification includes two sets of calculations representing two perspectives; one for the number of specific resources and one for the frequency of their use in the task. Data shows that the termino-lexicographic sub-category is the least represented, with one translator being classified as using termino-lexicographic resources in terms of their number, and three in terms of the frequency of use of specific resources. The mixed and the text-based spectrum are more densely populated, with an equal number of translators in each perspective - six for the number of specific resources and seven for the frequency.

<table>
<thead>
<tr>
<th>Type of dominance</th>
<th>Distribution</th>
<th>Number of specific resources used in the task</th>
<th>Frequency of specific resources used in the task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Termino-lexicographic</td>
<td>0%-33% text-based vs. 67%-100% termino-lexicographic</td>
<td>T5</td>
<td>T5, T8, T14</td>
</tr>
<tr>
<td>Mixed User</td>
<td>34%-66% text based vs. 34%-66% termino-lexicographic</td>
<td>T1, T3, T7, T8, T9, T14, T15</td>
<td>T1, T3, T7, T10, T13, T15</td>
</tr>
<tr>
<td>Text-based dominant</td>
<td>67%-100% text-based vs. 0% 33% termino-lexicographic</td>
<td>T4, T6, T10, T11, T12, T13, T16</td>
<td>T4, T6, T9, T11, T12, T16</td>
</tr>
</tbody>
</table>

Table 6.5 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NATURE OF RESOURCES USED

The most extreme use of text-based resources, mainly parallel texts, was displayed by T11, who accessed just three termino-lexical resources on four occasions but 24 text-based resources on 113 occasions. The most typical mixed user was T15, with an almost 50-50 ratio of using termino-lexicographic and text-based resources, both in terms of the number of specific resources and frequency of accessing those resources.
T15 used 6 specific termino-lexical resources 35 times and 8 text-based resources 42 times. The one translator classified as termino-lexicographic dominant, T5, accessed two dictionaries once each and Google search engine once.

With regard to the third type of resource discussed in this section, MT, it was mentioned earlier that it was an auxiliary category. Therefore, participants were classified as an MT user in addition to one of the other categories discussed above. Table 6.6 below shows that three individual translators (T5, T11 and T14) were classified as MT users, based on their use of MT in the search and discovery mode in the translation task. A total of four translators used machine translation, but of these, three used it as a resource for ad-hoc terminology research in a search and discovery mode rather than a translation tool applied to the text as a whole in a post-editing mode²⁰⁶. T5 used MT on eight occasions, T11 on twenty occasions and T14 used it once. Table 6.6 below shows a summary of the classification of individual translators according to their use of termino-lexicographic resources, text-based ones and MT. It shows how many translators and which particular ones were classified as termino-lexicographic dominant, text-based dominant, mixed user and MT user based on two perspectives: the number of specific resources used in the task and the frequency of use of these resources.

<table>
<thead>
<tr>
<th></th>
<th>Termino-lexicographic</th>
<th>Mixed user</th>
<th>Text-based</th>
<th>MT user</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of specific resources used</td>
<td>Frequency of specific resources use</td>
<td>Number of specific resources used</td>
<td>Frequency of specific resources use</td>
<td>Number of specific resources used</td>
</tr>
<tr>
<td>Number of translators</td>
<td>1 3</td>
<td>7 6</td>
<td>7 6</td>
<td>3 3</td>
</tr>
<tr>
<td>Individual translators</td>
<td>T5, T8, T14</td>
<td>T1, T3, T7, T8, T9, T14, T15</td>
<td>T4, T6, T10, T11, T12, T13, T16</td>
<td>T5, T11, T14</td>
</tr>
</tbody>
</table>

Table 6.6 SUMMARY OF CLASSIFICATION OF TRANSLATORS ACCORDING TO THE TYPE OF RESOURCES THEY USE

Compared to the findings of Désilets et al. (2009), whose study participants used slightly more termino-lexicographic resources than text-based ones²⁰⁷, the participants in the present study used more text-based resources. This confirms Désilets et al.’s (2009) conclusion that “corpus-based technologies have clearly made it into the mainstream and that translators are comfortable using them”, but stands in stark opposition to their statement that they “do not seem to have replaced the termino-lexicographic resources” (Désilets et al., 2009). However, the five-year gap between Désilets et al.’s study and the present study during which the text-based resources might have gained more momentum, has to be acknowledged.

²⁰⁶ T8 and T14 also used MT as applied to the whole document, i.e. in a post-editing mode.

²⁰⁷ Désilets et al. calculate the use of resources by considering the number of problems for which a particular type of resource was consulted and the total number of times it was consulted. The second calculation reflects the ‘frequency of use’ perspective in this study.
6.1.2. Research direction

*Research direction* contributes to the RTUT by adding a dimension associated with the orientation of the search towards the source or the target language, which, to some extent, determines the choice of the resource. The more ST-oriented the searches, the more likely termino-lexicographic resources will be used. This is because ST-oriented searches are often associated with finding definitions or equivalence seeking, based on a particular research unit found in the text. When it comes to testing a hypothesis in a target language, this is most likely to be carried out via a text-based resource as it is associated with ‘checking’ or ‘confirming’ actions, which can be done via Google search or within a parallel text. An example of such clear-cut separation of the two types of research direction associated with a certain type of resource is T15, who almost exclusively searched in termino-lexicographic resources for ST-oriented queries (highlighted in green) and in text-based resources for TT-oriented queries (highlighted in yellow) (see Table 6.7).

<table>
<thead>
<tr>
<th>Research Direction</th>
<th>Query</th>
<th>Specific Resource used</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>rise</td>
<td>WordReference</td>
</tr>
<tr>
<td>TT</td>
<td>monnaie électronique bitcoin</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>banque WIR</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>système HOURS Ithaca</td>
<td>Google</td>
</tr>
<tr>
<td>ST</td>
<td>commodities</td>
<td>WordReference</td>
</tr>
<tr>
<td>TT</td>
<td>monnaie basée sur les merchandises</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>étalon or</td>
<td>Google</td>
</tr>
<tr>
<td>ST</td>
<td>popular</td>
<td>WordReference</td>
</tr>
<tr>
<td>ST</td>
<td>portfolio</td>
<td>WordReference</td>
</tr>
<tr>
<td>ST</td>
<td>disruptive</td>
<td>WordReference</td>
</tr>
<tr>
<td>TT</td>
<td>perturbateur</td>
<td>Thesaurus.com</td>
</tr>
<tr>
<td>TT</td>
<td>subversif</td>
<td>Synonymes.com</td>
</tr>
<tr>
<td>TT</td>
<td>subversif</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>un système subversif</td>
<td>Google</td>
</tr>
<tr>
<td>ST</td>
<td>mainstream</td>
<td>WordReference</td>
</tr>
<tr>
<td>ST</td>
<td>fiat money</td>
<td>Google</td>
</tr>
<tr>
<td>ST</td>
<td>fiat money</td>
<td>WordReference</td>
</tr>
<tr>
<td>TT</td>
<td>monnaie LR de Liberty Reserve</td>
<td>Google</td>
</tr>
<tr>
<td>ST</td>
<td>hard currency</td>
<td>WordReference</td>
</tr>
<tr>
<td>TT</td>
<td>monnaie des espèces</td>
<td>Google</td>
</tr>
<tr>
<td>ST</td>
<td>euro capital letter or not</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>euro avec majuscule ou pas</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>euro avec majuscule ou pas</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>monnaie adossee a une banque</td>
<td>Google</td>
</tr>
<tr>
<td>ST</td>
<td>backed</td>
<td>WordReference</td>
</tr>
<tr>
<td>TT</td>
<td>monnaie garantie par la banque centrale</td>
<td>Google</td>
</tr>
<tr>
<td>TT</td>
<td>garantie par la banque centrale</td>
<td>Google</td>
</tr>
</tbody>
</table>

208 Out of the 865 instances of resource use by the sixteen participants in the translation task, not one instance of using a dictionary in the L2 → L1 direction, for example to confirm a hypothesis, was observed.
Research direction is indicated by typing a query into a resource, either in the source language or the target language. If a translator initiates a search in the source language, then the action is deemed ST-oriented. Conversely, if the search is initiated in the target language, then the action is considered TT-oriented. Therefore, the presence of the query is associated with the presence of the relevant code (ST-oriented or TT-oriented). The number of combined ST-oriented and TT-oriented actions does not equal the total number of instances of resource use. In other words, the codes referring to ST or TT-oriented actions were not used each time a resource was accessed. This is because queries were not entered each time a resource was used. Although each research episode was initiated with a query in either ST or TT, subsequent consultations in the same research episode may only have been associated with further
exploration of the same resources by the use of hyperlinks or with simply clicking on subsequent pages of the same resources, in which case no new query was present. However, whenever a query was modified in the same research episode, then another instance of ST or TT-oriented action was coded. Table 6.8 below illustrates an example of when an ST-oriented code was not assigned.

<table>
<thead>
<tr>
<th>T15</th>
<th>Research episode 5 (Commodities based currency)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step</td>
<td>WordReference bilingual dictionary WordReference discussion forum Google</td>
</tr>
<tr>
<td>Research query typed</td>
<td>commodities n/a monnaie basée sur les marchandises</td>
</tr>
<tr>
<td>Research direction code</td>
<td>ST-oriented n/a TT-oriented</td>
</tr>
<tr>
<td>Reason, why code was assigned/not, assigned</td>
<td>Translator entered query in the ST language (English) Translator accessed WordReference forum from the WordReference bilingual dictionary. Although a different resource was used, no query was entered Translator entered query in the TT language (French)</td>
</tr>
</tbody>
</table>

Table 6.8 EXAMPLE OF HOW RESEARCH DIRECTION CODES WERE ASSIGNED IN RESEARCH EPISODES

Figure 6.1 below shows the observed ST-oriented vs. TT-oriented searches as performed by individual translators during the translation task in absolute frequencies and percentages. The most extreme ST-oriented participant, T3, made 53 consultations in the source language, which amounted to 93% of all her searches. This coincided with heavy use of a bilingual dictionary, but she also made frequent ST-oriented queries in ProZ Glossaries and Google search engine. On the other end of the spectrum was T16 whose searches originating in the target language accounted for 69% of all searches; however, overall she performed 4.5 times fewer searches than T3.

Table 6.9 below illustrates how all individual participants were classified according to the distribution of their ST-oriented and TT-oriented searches in the translation task. The red end of the spectrum indicates ST-oriented searches and the blue indicates TT-oriented searches.

Figure 6.1 SOURCE TEXT ORIENTED SEARCHES VS. TARGET TEXT ORIENTED SEARCHES MEASURED IN ABSOLUTE FREQUENCIES AND IN PERCENTAGES

Table 6.9 EXAMPLE OF HOW ALL INDIVIDUAL PARTICIPANTS WERE CLASSIFIED ACCORDING TO DISTRIBUTION OF THEIR ST-ORIENTED AND TT-ORIENTED SEARCHES IN THE TRANSLATION TASK
The prevalent direction of searching in the translation task was by ST-oriented queries, which, in the case of this study, was English. A total of eight participants demonstrated ST-oriented search behaviour, followed by mixed behaviour exhibited by six translators. Only one participant (T16) displayed TT-oriented dominance. It was not possible to establish research direction preference for T2 who did not conduct any research. Table 6.10 below illustrates the classification of individual translators into ST-oriented, Mixed and TT-oriented sub-categories of research direction, showing which particular translators were classified into which category and the total number of translators representing each category.

<table>
<thead>
<tr>
<th>Individual translators</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST-oriented</td>
<td>T3, T5, T7, T8, T9, T10, T13, T14</td>
</tr>
<tr>
<td>Mixed</td>
<td>T1, T4, T6, T11, 12, T15</td>
</tr>
<tr>
<td>TT-oriented</td>
<td>T16</td>
</tr>
</tbody>
</table>

**Table 6.10 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE RESEARCH DIRECTION**

The results of this analysis pointing towards the fact that research was predominantly carried out in the source language are somewhat surprising. In the context of corpus-base translation studies, Bernardini (2011:2) claims that since Toury’s target-oriented approach to translation, an ST-oriented, equivalence-focussed approach has been viewed as “somehow backward-looking”. In translator training too, much emphasis has been put on using parallel texts to search for functional rather than formal equivalence and the ability to use the hypothesis-driven strategy (translation problem ➔ hypothesis in TL ➔ checking...
hypothesis in TL sources) is associated with the proficiency and experience of professional translators. In her study of two experienced translators\(^{209}\) and four students, Enríquez Raído (2014:84) found that students primarily consulted dictionaries and chose source-text keywords while professionals mainly searched in parallel texts using keywords in their target language. Similar findings were obtained by Prassl (2010) and Volanen (2015). The findings of the present study reveal that even within a group of professional translators there is a tendency to rely on searches in the source language for much of the research. This, however, must be seen in the context of the study design, which only allowed one source language, i.e. English.

When dominant research direction was compared to the length of experience of translators taking part in this study, no relationship was found between these two variables. In fact, three participants (T7, T13 and T14) with a very high proportion of ST-oriented searches (83%, 86% & 86% respectively) were found to be on the more experienced end of the spectrum, with T13 having over twenty one years of experience (see Table 6.11). Therefore, the results of this study undermine the commonly held view that research direction depends on the level of experience and that the more experienced the translators the more TT-oriented research they conduct.

<table>
<thead>
<tr>
<th>Translator</th>
<th>T11</th>
<th>T16</th>
<th>T15</th>
<th>T10</th>
<th>T9</th>
<th>T8</th>
<th>T4</th>
<th>T3</th>
<th>T14</th>
<th>T12</th>
<th>T7</th>
<th>T6</th>
<th>T13</th>
<th>T5</th>
<th>T1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of experience</td>
<td>3-5</td>
<td>6-10</td>
<td>6-10</td>
<td>6-10</td>
<td>6-10</td>
<td>6-10</td>
<td>6-10</td>
<td>11-20</td>
<td>11-20</td>
<td>11-20</td>
<td>11-20</td>
<td>21-30</td>
<td>31+</td>
<td>31+</td>
<td></td>
</tr>
<tr>
<td>Research direction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target language</td>
<td>Pl</td>
<td>Es</td>
<td>Fr</td>
<td>Es</td>
<td>Pl</td>
<td>Id</td>
<td>Es</td>
<td>Pl</td>
<td>Es</td>
<td>Es</td>
<td>Pt (Brazil)</td>
<td>Nl</td>
<td>Hu</td>
<td>Pt</td>
<td>Es</td>
</tr>
</tbody>
</table>

Table 6.11 EXPERIENCE VS. RESEARCH DIRECTION

Also, while it could be argued that familiarity with a domain might foster TT-oriented research, two translators who performed the most TT-oriented searches did not specialise in the business/economic/financial domains the text belonged to. They were not familiar with the terminology in the source text and yet they were able to formulate queries in the target language based on the parallel texts they consulted during the task.

In the light of these findings, it could be argued that directionality of research does not necessarily have to depend on experience or familiarity with the source-text domain alone, but can also be a matter of preference or habit, although further studies with multiple source languages are needed to reveal more patterns related to directionality. Some may argue that more fluent translations are produced as a result of the TT-oriented research, but ST-oriented research cannot be underestimated, especially when it comes to accuracy in researching a difficult subject matter. However, quality assessment is outside the scope of this research and further research is needed to establish whether research direction affects the quality of the translation.

---

\(^{209}\) One of the experienced translators had, in fact, only three years of casual experience in translation, and had a PhD in translation.
6.1.3. Research strategy

The term *research strategy* in the present work relates to different ways of initiating searches via the web. It distinguishes those translators who have established search routines and who initiate searches from their favourite resources from those who prefer using search engines as a strategic resource from which all other resources can be accessed. It would appear that those translators who have an organised system of tabs with their favourite resources or know which resource they want to use for which particular purpose tend to use more termo-lexicographic resources. Those who are less organised frequently rely on a search engine to locate the content, irrespective of which resource it might come from.

Many different classifications relating to the concept of search strategy have been proposed in information behaviour literature (Broder, 2002; D. Green, 2000; White & Iivonen, 2001) and also in literature from other areas, for example in tourism management (Chang & Caneday, 2011). In translation studies, references to search strategies are found in Austermühl (2001) and Enríquez Raído (2014). In Table 6.12 below, those classifications are mapped onto the three research strategies used in this study.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Navigational search</td>
<td>Direct address</td>
<td>Visiting known sites</td>
<td>Institutional searches</td>
<td></td>
<td></td>
<td>Bottom-up</td>
</tr>
<tr>
<td>Web directories</td>
<td>Subject directory</td>
<td>Subject directory</td>
<td>Thematic searches</td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>Informational search</td>
<td>Search engines</td>
<td>Search engine</td>
<td>Keyword search</td>
<td></td>
<td></td>
<td>Search Engine Queries Top-down</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Search engines</td>
<td></td>
<td></td>
<td>Browse searches</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Web browsing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transactional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.12 VARIOUS CLASSIFICATIONS OF TYPES OF SEARCH QUERIES

Although all the above classifications can be loosely mapped onto the categories of search strategies used in the present work, their coverage does not adequately reflect the research actions observed in the translation task of the Main Study. The widely accepted general typology of the types of search queries in information behaviour literature is the classification of Broder (2002), which covers three types of search
queries based on user intent: informational, navigational and transactional\textsuperscript{210}. Broder’s informational task refers to finding an answer irrespective of the location as long as it is reliable. It can be compared to the top-down strategy, where a search engine is used to locate information. Navigational tasks refer to searches in which users have a resource in mind. These can be compared to the bottom-up strategies. The transactional category refers to the intent of making a “web-mediated activity” such as shopping or downloading files (Broder, 2002:3) and therefore, while the navigational and informational search techniques can be loosely mapped onto the top-down and bottom-up strategies used in this study, the translational task cannot be compared to any behaviour of a translator found in the present study.

The other proposed classifications also do not fully reflect the translators’ search strategies that were observed in the present study. The reasons are twofold. Firstly, all the above classifications apart from Broder’s, include a category of web directories, also referred to as subject directories, which are defined by Green (2000:125) as a selective, pre-defined list of websites compiled by human editors and categorised according to subject/topic. Green claims that “web directories are popular with Internet users looking for particular information because they feel that they have a head start in identifying ‘the best of the web’ for the topic that they are interested in” (Green, 2000:125). Furthermore, although Austermühl (2001:55) lists thematic searches via subject trees (directories) as a viable option for translators, in this study this behaviour was not observed. This category was also missing in Enríquez Raído’s observations (2014:123).

Secondly, in the present study an additional search strategy was observed which can be described as a combination of the top-down and bottom-up technique (see Figure 6.2). It was used by only one participant, T9, but it was used extensively, as shown in Figure 6.3 and Figure 6.4.

This type of strategy is what I have called a hybrid strategy in which keywords are inserted into a search engine, but the way the query is formulated reveals the translator’s intention of using a particular resource in a particular language. Therefore, the top-down technique is used as a vehicle for executing a bottom-up strategy. Examples of T9’s queries are: ‘falter megasłownik’, ‘portfolio ProZ polish’, ‘longevity megasłownik’, ‘longevity proz polish’, ‘established proz polish’ or ‘haircut multilingual dictionary’. What is also interesting is that in most cases the consultation is restricted to reading the ‘snippets’, i.e. the Google result pages without actually clicking into a resource. This raises a methodological question as to whether the resources mentioned in the search query were actually consulted in these cases or whether it was the Google search engine itself. For the purpose of this study and in line with the definition of research step (see section 4.2.) it was deemed appropriate not to count specific resources that were mentioned in a search query (e.g. ProZ or Megasłownik) as consulted unless they were physically accessed. In such cases, a search engine was considered to be the resource used.

\textsuperscript{210}Broder’s typology is based on the user’s various intents, i.e. “the type of resource desired in the user’s expression to the system” (Jansen et al., 2008:1252).
Taking into account these two reasons, a classification of search strategies that reflects the lack of the observed web directory searches and the presence of an additional hybrid type was adopted. In the present research, three search strategies were identified from the data: top-down strategy in which the searcher initiates the query from a search engine, bottom-up, in which the searcher initiates the query from a known source and a hybrid strategy in which these two approaches are combined.

Figure 6.2 T9’s USE OF HYBRID RESEARCH STRATEGY

Figure 6.3 and Figure 6.4 below show, both in terms of relative and absolute frequencies, that translators used a mix of top-down and bottom-up strategies in various proportions, and one translator also used the hybrid strategy. A total of ten translators used more bottom-up strategies than top-down strategies and five used more top-down techniques than bottom-up ones.

Figure 6.3 RESEARCH STRATEGIES AS USED BY INDIVIDUAL TRANSLATORS MEASURED IN PERCENTAGES

186
The present data shows that bottom-up strategies were recorded 317 times and accounted for 63% of all recorded search actions, while top-down strategies were observed 172 times and accounted for 34% of observed actions. The hybrid type was used 18 times and accounted for 3% of the search-initiating actions. However, it is worth recalling that not all instances of resource use were associated with one of the categories, as subsequent clicking on further pages of the same resource were not included in the calculations. Therefore, out of the 865 instances of resources use in the whole task by all participants, only 507 were assigned to one of the three search strategy categories.

From Figure 6.3 and Figure 6.4 it can be seen that the participants used more bottom-up techniques than top-down techniques, with two thirds of all actions being bottom-up. This would suggest that translators preferred to start searching with a particular resource in mind. Whether the resource was optimal to solve a particular translation problem and whether the search was successful is a matter that was beyond the scope of this study as neither the quality of the translation nor the success rate of the searches were assessed.

In Enríquez Raído’s study, direct address searches, i.e. bottom-up, accounted for almost 75% in one task and 50% in the second task (Enríquez Raído, 2014:155). However, only 59 and 54 overall actions, respectively, were carried out by her six participants, compared to 319 in the present study. Additionally, Enríquez Raído based her findings on calculations of only initial actions, i.e. the first resource to be consulted in what is called here a research episode, while the above Figure 6.3 and Figure 6.4 included all recorded actions within the research episode, but excluded clicks on further pages of the same resource. However, in order to make a like for like comparison, a calculation was performed to include only initial actions taken in the translation task of the present study showing that that the proportions have remained very similar, regardless of whether only initial actions or all recorded actions were taken into account (see Figure 6.5).
Therefore, it can be confidently concluded that bottom-up search strategies accounted for two thirds of search actions. However, unlike in Enríquez Raído’s study (2011), most participants in this study were observed to use both types of strategies. Enríquez Raído reports that in her Task 1, participants showed “little or no variation of initial search action […] i.e. the participants with the least translation (and search) experience started their searches from ‘the known,’ while the most experienced translator (and searcher) preferred to initiate his searches with ‘the unknown’” (Enríquez Raído, 2011:474-475). This improved in Task 2, where it is reported that “the students’ choice of initial search actions was wider than in Task 1, and involved a combination of direct address searches (to access reference sites) and search engine queries”. Furthermore, while all but one of Enríquez Raído’s participants “chose the same source of information regardless of the type of question asked” (Enríquez Raído, 2014:124), participants in this study were largely found to apply search strategies according to the type of knowledge sought to satisfy their research needs.

The detailed distribution of the searches according to top-down or bottom-up strategy is illustrated in Table 6.13 below. The red end of the spectrum indicates bottom-up searches and the blue indicates top-down ones. The hybrid search strategy is marked by the colour purple. Eight translators displayed bottom-up dominance, four participants used mixed strategies and three showed top-down dominance. While two translators displayed the extreme use of bottom-up strategies (i.e. used this type of strategy in over 90% of their searches), such extreme use of top-down strategies was not observed, with the highest user of top-down strategies (T4) performing 77% of her searches in a top-down manner.
Table 6.13 DISTRIBUTION OF BOTTOM-UP AND TOP-DOWN SEARCHES

Table 6.14 below illustrates the classification of individual translators into Bottom-up, Mixed, Top-down and Hybrid sub-categories of research strategy showing which particular translators were classified into which category and the total number of translators representing each category.

The dominance of the bottom-up approach in research strategies indicates that the strategy of accessing resources from a known resource is currently preferred by translators. An explanation can be derived from a web search study of 23 subjects using eye-tracking carried out by Lorigo et al. (2006:1130) which concludes that informational tasks took longer than navigational ones, which, in the context of translation research, would mean that bottom-up techniques take less time and, therefore, can be seen as more effective, assuming an appropriate solution was found. However, with the increased number of available resources, the challenge lies in identifying the most appropriate resources to use for each translation.
problem. As one participant at the METM Meeting in 2013\textsuperscript{211} noted, with so many resources to choose from, performing a search ensuring wider recall using a search engine can sometimes be more effective.

6.2. The Resource Type User Typology (RTUT)

As shown in section 6.1., participants in this study displayed varied patterns of behaviour in relation to their choice of online resources for their translation-oriented research. Three different categories were used to systematise those patterns: the nature of resources, research direction and research strategy. These categories were used to classify individual translators into sub-groups according to their observed use of resources in terms of each of these three aspects. The outcomes of these classifications will be used in this section to build profiles for the Resource Type User Typology. These profiles are indicative of the types of resources that are preferred by participating translators and of how they go about accessing them during the translation process. For visualisation purposes, the categories are colour coded and run on a colour spectrum from red to blue (see Table 6.15). The red spectrum indicates the use of resources that are of a termino-lexicographic nature as well as ST-oriented searches and bottom-up techniques that are mostly associated with the use of these types of resources. The other end of the spectrum is coloured blue and indicates resources that are text-based, as well as TT-oriented searches and top-down strategies, mostly associated with these types of resources. In what follows, these three categories feeding into the Resource Type User Typology will be integrated, showing how the characteristic behaviours of the individual translators have been moulded into the following RTUT profiles: The Dictionary Enthusiast, The Mixed Type, The Parallel Text Fan and the MT Adopter, the latter being a profile that is supplementary to the other three.

<table>
<thead>
<tr>
<th>Nature of resource</th>
<th>Dictionary Enthusiast</th>
<th>Mixed User</th>
<th>Parallel Text Fan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research direction</td>
<td>ST-oriented</td>
<td>Mixed</td>
<td>Corpus-based</td>
</tr>
<tr>
<td>Research strategy</td>
<td>Bottom-up</td>
<td>Mixed</td>
<td>TT-oriented</td>
</tr>
</tbody>
</table>

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|}
\hline
Nature of resource & Dictionary Enthusiast & Mixed User & Parallel Text Fan \\
\hline
Research direction & ST-oriented & Mixed & Corpus-based \\
\hline
Research strategy & Bottom-up & Mixed & TT-oriented \\
\hline
\end{tabular}
\caption{COLOUR CODING OF CATEGORIES IN RTUT}
\end{table}

Table 6.16 below contains a heat map showing all three categories (the nature of resources being composed of two aspects - the number of resources and the frequency of their use) indicating each individual translator’s previously assigned classification. The three profiles (RTUT types) are assigned by calculating weights added to each of the categories, with each cell of the red spectrum being given a value of 3, each cell of a mixed colour having a value of 0 and each cell in the blue spectrum with a value of -3. The values of the weights are arbitrary and were assigned in order to create a numerical range of the spectrum on which each participant could be placed. The colour spectrum helps to visualise the behaviours; clustering around the red end of the spectrum indicating the Dictionary Enthusiast profile, around the blue end of the spectrum indicating the Parallel Text Fan or having a mixture of colours which represents the Mixed Type. As mentioned before, the MT user is an additional category and is represented by the colour purple.

\textsuperscript{211} This refers to a discussion that took place after the presentation of some preliminary findings of this research at the METM Meeting in Tarragona, Spain in October 2013.
The Resource Type User Classification is presented below, showing the four types of resource user, together with their typical resource use behaviours as observed in the translation task of the Main Study.

The Dictionary Enthusiast prefers to consult termino-lexicographic resources such as dictionaries, term banks or glossaries. It could be someone who uses a variety of different dictionaries, glossaries etc., but not necessarily very often or someone who uses just one dictionary, but very frequently. Dictionary Enthusiasts will often have their favourite resources bookmarked or otherwise organised for quick access. They will often initiate searches in their source language and because of the extensive use of the resources they know well, the preferred way of accessing them is bottom-up.

The Parallel Text Fan is the opposite of the Dictionary Enthusiast. Their preferred modus operandi for accessing resources is via a keyword in a search engine, i.e. top-down, and they will often go several pages deep in their search. They will not normally have an organised system of resources and will often use the search engine to find them. T4, for example, typed a resource she had in mind into Google each time she wanted to access it. They will often perform extensive TT-oriented research and will do so using parallel texts in the form of web pages and online documents. T11, for example, used eighteen different websites to conduct the parallel text research not only to understand more about the subject but also to find equivalents.

The Mixed Type shares the characteristics of both the Dictionary Enthusiast and the Parallel Text Fan and the MT Adopter can be any of these types who uses machine translation in their work.
As mentioned before, the Resource Type User Typology offers a systematic approach to classifying translators according to their preferred type of resources and the different ways these resources are accessed and used. Although like any other attempts to classify human behaviour this typology can only be treated as a general direction and approximation rather than as an exact depiction of the behaviour of each of the participating translators, it does offer an attempt to develop profiles of resource use in translation studies research.

The use of online resources by professional translators is a complex issue and one that is subject to evolution. Translators’ online behaviour can change as a result of the emergence of new technologies and resources, and the rate of their adoption by the profession. This is why the Resource Type User Typology is embedded in its own time and it may be subject to change as language technology advances. Before search engines came to be “amongst the most popular destination sites on the Web” (Green, 2000:125) and thus indispensable for translators, The Dictionary Enthusiast would have been the most dominant resource type user, if not the only one. This is because paper dictionaries, hand-made glossaries and card index systems were the primary sources of terminological information, and bottom-up techniques would have been widely used by professional translators as search engines were still in their nascent state. The emergence of the Parallel Text Fan could, therefore, be attributed to the fact that the written legacy of humankind has been digitalised, published on the Internet and made available via search engines. The influence of technology on the changes in search procedures can even be noticed in the absence of thematic searches (i.e. searching via subject trees such as Yahoo directories) in the present study. This was referred to by Austermühl as early as 2001 and also mentioned by Enríquez Raído (2011). As subject directories have become supplanted by search engines with their ever increasing information retrieval capabilities,
accessing information via subject trees, as mentioned before, has become obsolete. Many such changes, driven by the developments in the semantic web, information retrieval, artificial intelligence and related areas, will be shaping the resource type user in the future. It is therefore of interest to translation process research and for the development of translation tools and resources to monitor the patterns of interaction with online resources by professional translators and to see, as Zapata (2015:151) suggests, “how the interaction with […] information tools can change over time”.

Despite the fact that the procedural aspects of translators’ interactions with online resources might be dependent on technological developments, human factors of affect and cognition still need to be taken into account as certain aspects of translators’ work do not necessarily lend themselves to change. While technology might influence translators’ choices of which resources to use, it might not entirely change the way they approach their translation research. While in future some translators might rely wholly on search engines to find the information they need in a top-down manner (especially as the search engines become more personalised and learn from user behaviour), others will still prefer to have a library of tabs in which they store their favourite resources and access them using the bottom-up strategy. By the same token, while some translators will continue to use search engines in an exploratory way, others will still require a pre-designed method of using the web resources. These are the more constant aspects of translation-oriented research relating to personal style which will be explored in the next Chapter.
7. Individual differences - Typology of Translator Research Styles (TTRS)

This chapter will concentrate on individual differences relating to research style in the online research behaviour of professional translators as observed in the translation task of the Main Study. In contrast to the previous chapter, where the individual differences more susceptible to technological changes were discussed (i.e. resource type user), the present chapter is concerned with a more constant aspect of the way translators engage with reference material, i.e. their information behaviour. While changing technology might influence the translator’s choice of resources for conducting research, as noted at the end of Chapter 6, it might not entirely change their approach to research, i.e. their research style. The way research is conducted and, consequently, the resources consulted, can be perceived as a mixture of various factors that are independent of the type of resource or the search techniques used. As mentioned in section 2.6., other important ingredients of this mixture, such as personality traits, can also come into play here; however, it is beyond the scope of this thesis to delve into aspects of research behaviour other than from an information behaviour perspective. That said, some of the aspects analysed, especially some of the secondary, qualitative categories, could be seen as related to personality. In what follows, the various factors related to translators’ online research behaviour will be explored in detail, leading to the formulation of the Typology of Translator Research Styles.

7.1. Categories feeding into the Typology of Translator Research Styles

Two groups of categories of translator research behaviour, primary and secondary, are considered here in order to create the Typology of Translator Research Styles. The primary categories originate from numerical data and were calculated for all the participants based on their screen recordings. The numerical data such as the number of research units, the number of specific resources accessed or the time spent on research etc. provide the basis for the formation of categories which represent ranges of behaviour that can be represented on a Likert scale. Some of the primary categories are compound, i.e. composed of two or more sub-categories. The six primary categories and their sub-categories are divided into two groups to represent volume-related and time-related features (See Table 7.1 below).

<table>
<thead>
<tr>
<th>Volume-related</th>
<th>Primary Categories</th>
<th>Sub-categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Research Unit (RU) Volume</td>
<td>Number of RUs researched in the task</td>
</tr>
<tr>
<td>2</td>
<td>Resource Volume</td>
<td>Number of specific resources used in the task</td>
</tr>
<tr>
<td>3</td>
<td>Research Intensity</td>
<td>The number of steps per RU (mean)</td>
</tr>
<tr>
<td>4</td>
<td>Research Diversity</td>
<td>The number of steps per RU (weighted mean)</td>
</tr>
<tr>
<td>5</td>
<td>Research Time</td>
<td>Number of types of resources used in the task</td>
</tr>
<tr>
<td>6</td>
<td>Research Pace</td>
<td>Average duration of research steps</td>
</tr>
<tr>
<td>Time-related</td>
<td></td>
<td>Research time (hrs:min:sec)</td>
</tr>
</tbody>
</table>

Table 7.1 THE PRIMARY TTRS CATEGORIES AND THEIR SUB-CATEGORIES
The *secondary categories* are of a qualitative nature and were gathered from the screen recordings, profile questionnaires and post-task e-mail questionnaires. These have been coded on an *ad hoc* basis and have only been applied to those participants who exhibited behaviours described in those categories. The following fourteen secondary categories were identified from the data:

<table>
<thead>
<tr>
<th>Secondary Categories</th>
<th>Short description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Deep searches</td>
<td>Deep engagement with resources characterised by longer periods of consultation or performing a chain of online interactions, often modifying queries and changing the type of resource</td>
</tr>
<tr>
<td>2 Shallow searches</td>
<td>Shallow engagement with resources characterised by one-step consultations, quick checks in termino-lexical resources or a rapid Google search with snippet viewing</td>
</tr>
<tr>
<td>3 Meandering research path</td>
<td>Diverging from a straight path by following internal links and hyperlinks of a resource</td>
</tr>
<tr>
<td>4 Parallel research sequence</td>
<td>Consulting more than one resource at the same time by opening several windows in a parallel fashion</td>
</tr>
<tr>
<td>5 Planning</td>
<td>Considering what research activities will be required for the task</td>
</tr>
<tr>
<td>6 Squirreling</td>
<td>Saving information for future retrieval</td>
</tr>
<tr>
<td>7 Snippet viewing</td>
<td>Consulting the search engine results without following the links</td>
</tr>
<tr>
<td>8 Using advanced queries</td>
<td>Using advanced procedural knowledge of web searching techniques</td>
</tr>
<tr>
<td>9 Repetitive behaviour</td>
<td>Repeating an initial search action consistently using the same resources in the majority of research episodes</td>
</tr>
<tr>
<td>10 Strategic behaviour</td>
<td>The presence of a clear and consistent ‘tactic’ for finding the required information with regard to word and world knowledge.</td>
</tr>
<tr>
<td>11 Drive for perfection</td>
<td>Exerting significant amount of effort to meet one’s research need</td>
</tr>
<tr>
<td>12 Affective assessment</td>
<td>Engaging feelings or attitudes as part of the evaluation of a solution</td>
</tr>
<tr>
<td>13 High information retention</td>
<td>Self-declared level of retaining information</td>
</tr>
<tr>
<td>14 Attitude towards technology</td>
<td>Self-declared classification with regard to adopting new technologies</td>
</tr>
</tbody>
</table>

Table 7.2 THE SECONDARY TTRS CATEGORIES

Most of the categories discussed in this chapter have been developed for the purpose of this research and are not based on previous literature. The only exceptions are ‘squirreling’ and ‘attitude towards technology’ categories, the origins of which are discussed in the relevant sections. Some categories have been used in other studies under different names or in different contexts, e.g. ‘snippet viewing’ (Enríquez Raído, 2011; Volanen, 2015) under the heading SERPs (search engine results pages) or behaviour relating to ‘strategies’ (Asadi & Séguinot, 2005) used to describe translation styles (prospective thinking, translation on-screen).
7.1.1. Primary categories

As mentioned above, the primary categories of the TTRS which characterise translators’ research behaviour (as opposed to resource use behaviour) constitute the fundamental part of the typology. They are based on a quantitative analysis and refer to the many aspects of research behaviour listed in Table 7.1. As mentioned earlier, some of these categories are compound and each component received a separate calculation. Each of the sub-categories was divided into High, Medium and Low values based on the range of numerical values that were observed for each calculation, capturing the lower end of the spectrum, the middle range and the high end. Similarly to the Resource Type User Typology presented in the previous chapter, the categories are colour-coded and run from light pink to dark pink for easier visualisation. The light pink end of the spectrum refers to the activities on the lower side of the quantitative scale indicating low quantities and short timings while at the other end of the spectrum, the dark pink colour indicates high quantities and long timings.

<table>
<thead>
<tr>
<th>Primary Category</th>
<th>Sub-category 1</th>
<th>Sub-category 2</th>
<th>Sub-category 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research unit Volume</td>
<td>RU Volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource Volume</td>
<td>Number of specific resources</td>
<td>Frequency of use of specific resources</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Research Intensity</td>
<td>Number of steps per RU (mean)</td>
<td>Number of steps per RU (weighted mean)</td>
<td>Number of steps per RU (weighted median)</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Research Diversity</td>
<td>Number of types of resources</td>
<td>Number of specific resources</td>
<td>Number of specific resources per RU</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Research Time</td>
<td>Research time (hrs:min:sec)</td>
<td>Research time (%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Research Pace</td>
<td>Average duration of research step</td>
<td>Average duration of research episode</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.3 OVERVIEW OF PRIMARY CATEGORIES AND THEIR COMPONENTS

The values in each sub-category of a primary category are given descriptive labels corresponding to the Low, Medium or High values, thus affording a qualitative description of corresponding characterising behaviours for each of the categories as indicated in Table 7.4. These behaviours were labelled as Generous, Moderate and Frugal for the volume-related categories and Relaxed, Moderate and Rapid for the time-related categories (see Table 7.4).

---

212 The only exception is the research pace category, where low (short) duration of research steps and episodes equals high research pace.
The range of characterising behaviours

<table>
<thead>
<tr>
<th>Volume</th>
<th>Primary Categories</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RU Volume</td>
<td>Frugal</td>
<td>Moderate</td>
<td>Generous</td>
</tr>
<tr>
<td>2</td>
<td>Resource Volume</td>
<td>Frugal</td>
<td>Moderate</td>
<td>Generous</td>
</tr>
<tr>
<td>3</td>
<td>Research Intensity</td>
<td>Frugal</td>
<td>Moderate</td>
<td>Generous</td>
</tr>
<tr>
<td>4</td>
<td>Research Diversity</td>
<td>Frugal</td>
<td>Moderate</td>
<td>Generous</td>
</tr>
<tr>
<td>Time</td>
<td>5</td>
<td>Research Time</td>
<td>Rapid</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Research Pace</td>
<td>Rapid</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Table 7.4 COLOUR CODING AND CHARACTERISING BEHAVIOURS OF THE CATEGORIES FOR THE TYPOLOGY OF TRANSLATOR RESEARCH STYLES

In what follows, each primary category will be discussed, providing a calculation for each of its components with classification into Low, Medium and High values, as well as a final classification using the characterising behaviour labels Frugal, Moderate and Generous for primary categories 1-4 and Rapid, Moderate and Relaxed for primary categories 5 & 6.

7.1.1.1. Research unit (RU) volume

The first primary category to be considered for the Typology of Translator Research Style is the research unit volume (RU volume), which indicates the number of words, terms or phrases that were researched by each individual translator during the translation task. It differentiates translators who have high research needs and consult resources for every instance of doubt or uncertainty about an aspect of the text they are translating, from those who only pay attention to the key items they deem necessary to achieve the required quality of the translated text.

A total of 88 RUs was observed in all 16 translation sessions (see Appendix 10), but translators varied markedly with regard to the number of units they chose to research. The lowest observed number of research units was 0 the highest 43. Participating translators were divided into tertiles based on the number of units they researched. Table 7.5 below shows the number of research units per individual translator, together with the classification as High, Medium or Low volume of RUs. Half of the sample had Low RU volume, 25% were classified as Medium and the remaining three translators were classified as having High RU volume.
The individual translators are classified for the purpose of TTRS by assigning the following behaviour characterising labels to their RU volume: Generous for high RU volume, Moderate for medium RU volume and Frugal for low RU volume (see Table 7.6 below).

Table 7.6 RU VOLUME CLASSIFICATION

As noted before, establishing the nature of research units and the presence of outliers posed methodological challenges and are discussed in section 3.4.

7.1.1.2. Resource volume

*Resource volume* indicates the number of resources accessed by each individual translator during the translation task. It is one of the most important categories characterising the research style of a translator as it shows a predisposition towards using certain quantities of resources in their work. While some translators keep accessing various resources until they find what they consider a perfect (or near perfect) solution, others will accept the first viable solution and not pursue other possibilities. These differences are reflected in the volume of resources used to research these units.

Two calculations were performed to establish resource volume: the total *number* of specific resources accessed during the task and the *frequency* of their consultation in the translation task. It was mentioned in section 5.2.5. that while some translators may have few resources in their tool box, they may use them frequently. Conversely, while others may use many resources, they may use them less often. Therefore, the resource volume category takes into account not only the number of specific resources individual translators accessed during the task but also the total number of individual resource consultations, i.e. the frequency of use of these resources.

Regarding the total *number* of specific resources, as discussed in 5.2.4., translators participating in the task differed considerably with regard to the number of resources accessed. At one end of the spectrum, T2
showed outlier behaviour as he did not consult any resources, but in the post-task questionnaire he did say that he was very familiar with the subject of digital currency and that he did not need to research anything. However, when he does conduct research, he reports (in the profile questionnaire) that he normally uses one to three resources, mainly Wikipedia, Linguee and his own glossaries. On the other end of the spectrum, T11 consulted as many as 28 resources, of which 18 were parallel texts. These differences were captured by classifying translators into tertiles representing Low, Medium or High number of specific resources use.

A total of nine translators (56%) were classified as having Low number of resources, five as Medium and two as High. It must be said, however, that T11 used a very high number of web pages \((n=18)\) and, as each website was classified as a separate resource, this distorted the distribution scale, thus classifying T3 with 9 specific resources as Low\(^{213}\). Table 7.7 below shows how each individual translator was classified according to the number of specific resources they consulted during the translation task.

<table>
<thead>
<tr>
<th>Resource Volume (number of specific resources)</th>
<th>Translator</th>
<th>Number of specific resources</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>20</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>0</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>9</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T4</td>
<td>11</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>T5</td>
<td>4</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T6</td>
<td>7</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T7</td>
<td>5</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T8</td>
<td>5</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T9</td>
<td>14</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>T10</td>
<td>15</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>T11</td>
<td>28</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>T12</td>
<td>8</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T13</td>
<td>5</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T14</td>
<td>7</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>T15</td>
<td>14</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>T16</td>
<td>11</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Range: 28 (0-28)

0 - 10 Low

11 - 19 Medium

20 - 28 High

Table 7.7 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NUMBER OF SPECIFIC RESOURCES USED IN THE TRANSLATION TASK

As for frequency of specific resources use, as discussed in 5.2.5., translators participating in this study accessed specific resources between 0 (T2) and 138 (T11) times during the translation task. Based on the number of times translators accessed various resources, they were assigned one of the three values: High,

\(^{213}\) It could be argued that because the average number of specific resources used in a translation task was found to be 10, 9 resources in translator’s toolbox would normally classify as Medium, not Low. However, because conditions for the exclusion of outliers (2.5 z-scores from the mean) did not apply to this distribution, the ranges have not been adjusted to exclude T2 from the calculation.

199
Medium or Low. The range of frequency of resources use was surprisingly wide within the group. Even ignoring T2’s outlier behaviour, the range between 12 (T5) and 138 (T11) is still considerable, taking into account the fact that the text was 412 words long. A total of ten translators, which is more than half of the sample (62%), was classified as having Low frequency of use of specific resources and three were classified into each of the Medium and High categories. Table 7.8 below illustrates the frequencies of specific resources use and how individual translators were classified according to those frequencies.

<table>
<thead>
<tr>
<th>Translator</th>
<th>Frequency of specific resources use</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>100</td>
<td>High</td>
</tr>
<tr>
<td>T2</td>
<td>0</td>
<td>Low</td>
</tr>
<tr>
<td>T3</td>
<td>98</td>
<td>High</td>
</tr>
<tr>
<td>T4</td>
<td>36</td>
<td>Low</td>
</tr>
<tr>
<td>T5</td>
<td>12</td>
<td>Low</td>
</tr>
<tr>
<td>T6</td>
<td>32</td>
<td>Low</td>
</tr>
<tr>
<td>T7</td>
<td>32</td>
<td>Low</td>
</tr>
<tr>
<td>T8</td>
<td>37</td>
<td>Low</td>
</tr>
<tr>
<td>T9</td>
<td>87</td>
<td>Medium</td>
</tr>
<tr>
<td>T10</td>
<td>80</td>
<td>Medium</td>
</tr>
<tr>
<td>T11</td>
<td>138</td>
<td>High</td>
</tr>
<tr>
<td>T12</td>
<td>29</td>
<td>Low</td>
</tr>
<tr>
<td>T13</td>
<td>40</td>
<td>Low</td>
</tr>
<tr>
<td>T14</td>
<td>36</td>
<td>Low</td>
</tr>
<tr>
<td>T15</td>
<td>87</td>
<td>Medium</td>
</tr>
<tr>
<td>T16</td>
<td>25</td>
<td>Low</td>
</tr>
</tbody>
</table>

Range: 138 (0-138)
0-46 Low
47-92 Medium
93-138 High

Table 7.8 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE FREQUENCY OF USE OF SPECIFIC RESOURCES IN THE TRANSLATION TASK

The two components of resource volume discussed are brought together in Table 7.9 to show their combined effect on classifying individual translators as belonging to the General, Moderate or Frugal categories of research volume. Most of the sample, that is, ten participants, was characterised as Frugal (T2, T4, T5, T6, T7, T8, T12, T13, T14 & T16), four translators were labelled as Moderate (T3, T9, T10 & T15) and two (T1 & T11) were described as having Generous resource volume.

In certain cases where the classification in the two sub-categories differed, a judgement had to be made about the dominant pattern. In the case of T3, a Moderate category was assigned as a resultant mid-point value between the Low classification in terms of the number of specific resources and the High classification in terms of the frequency of resources use. For T4 and T16, the Frugal label was assigned as a compromise between a Medium and Low category for the number of specific resources and frequency of specific resources use respectively. The Low (Frugal) category was judged to be more dominant because it had a lower standard deviation within its range than the Medium category.
7.1.1.3. **Research intensity**

*Research intensity* communicates an important aspect of a translator’s research activity which relates to the number of consultations per research unit showing how much resource volume they are prepared to generate per research unit. It was observed that while some translators’ research is characterised by short, 1-2 step research episodes, others engage in lengthier pursuit of an acceptable solution, with the longest observed episode being 22 steps. Although the average number of steps per research episode for all translators was 2.8, they varied substantially with regard to the composition of their research episodes, which means that they differed in the number of resources they accessed to solve a particular problem.

The example below in Table 7.10 shows the variation between individual translators in the number of steps taken to satisfy their research need for the term ‘fiat money’. Although the average number of steps in this case was 3.7, the range was wide, between 1 and 8 steps.

![Table 7.9 RESEARCH VOLUME CLASSIFICATION](image)

*Table 7.9 RESEARCH VOLUME CLASSIFICATION*

<table>
<thead>
<tr>
<th>Resource Volume category label</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of specific resources</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Frequency of specific resources</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

**Table 7.10 THE NUMBER OF STEPS TAKEN BY INDIVIDUAL TRANSLATORS TO SATISFY RESEARCH NEEDS RELATED TO THE TERM ‘FIAT MONEY’**

<table>
<thead>
<tr>
<th>Log</th>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
<th>Step 5</th>
<th>Step 6</th>
<th>Step 7</th>
<th>Step 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>search engine</td>
<td>glossary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>search engine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td>MT</td>
<td>search engine</td>
<td>website</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6</td>
<td>termbank</td>
<td>search engine</td>
<td>knowledge based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T7</td>
<td>termbank</td>
<td>search engine</td>
<td>knowledge based</td>
<td>knowledge based</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T8</td>
<td>search engine</td>
<td>glossary</td>
<td>glossary</td>
<td>glossary</td>
<td>search engine</td>
<td>search engine</td>
<td>search engine</td>
<td></td>
</tr>
<tr>
<td>T9</td>
<td>search engine</td>
<td>knowledge based</td>
<td>knowledge based</td>
<td>search engine</td>
<td>knowledge based</td>
<td>search engine</td>
<td>search engine</td>
<td></td>
</tr>
<tr>
<td>T10</td>
<td>concordancer</td>
<td>knowledge based</td>
<td>knowledge based</td>
<td>knowledge based</td>
<td>search engine</td>
<td>knowledge based</td>
<td>search engine</td>
<td>knowledge based</td>
</tr>
<tr>
<td>T11</td>
<td>MT</td>
<td>dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T12</td>
<td>search engine</td>
<td>knowledge based</td>
<td>search engine</td>
<td>knowledge based</td>
<td>search engine</td>
<td>knowledge based</td>
<td>search engine</td>
<td></td>
</tr>
<tr>
<td>T13</td>
<td>dictionary</td>
<td>website</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T14</td>
<td>dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T15</td>
<td>search engine</td>
<td>dictionary</td>
<td>dictionary</td>
<td>dictionary</td>
<td>discussion forum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T16</td>
<td>concordancer</td>
<td>dictionary</td>
<td>Forum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| mean | 3.7 |        |        |        |        |        |        |        |

These differences can be observed in more detail when comparing these patterns between translators across all their research episodes. For example, Table 7.11 shows the difference in the number of steps between T13 and T16. The second column for each translator (steps/episode) shows that whilst most of T13’s research episodes are one-step consultations (14 out of 20), half of T16’s research episodes are four-step consultations, thus making T16 more intense than T13.
However, in order to classify translators into groups according to Low, Medium or High number of steps per research unit, this varying composition of research episodes had to be described by means of a central tendency, which had to reflect not only the average number of steps per research unit, but also the differences in the composition of research episodes of individual translators as illustrated in Table 7.11.

Therefore, three statistical methods of establishing an average value were used in a parallel fashion. Mean reflects a straight average where the total number of steps is divided by the total number of research episodes. Weighted mean takes into account the frequency of occurrence of research episodes of certain length by assigning a weight (that is, influence) corresponding to the number of times such research episode occurred. In other words, the research episodes that occurred more frequently (in terms of the length, i.e. the number of steps) carry more weight (i.e. importance). In the calculation all the steps in each research episode are multiplied by the assigned weight and added up, and then divided by the sum of all the weights, resulting in a more 'proportional' representation of the number of steps in research episodes. Median reflects the numbers of steps that fell in the midpoint of a frequency distribution and therefore also provide a useful insight. Due to the nature of the data, Mode was deemed unsuitable as a method of establishing a central tendency in this particular instance.

Table 7.12 below shows the average values for the number of steps per research episode obtained by using these three methods, together with the classification into Low, Medium or High number of steps per RU for each of the calculations.
<table>
<thead>
<tr>
<th></th>
<th>1 (mean)</th>
<th></th>
<th></th>
<th>2 (weighted mean)</th>
<th></th>
<th></th>
<th>3 (median)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of steps per RU</td>
<td>Classification</td>
<td>Number of steps per RU</td>
<td>Classification</td>
<td>Number of steps per RU</td>
<td>Classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>2.9</td>
<td>Medium</td>
<td>1.8</td>
<td>Low</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>0.0</td>
<td>n/a</td>
<td>0.0</td>
<td>n/a</td>
<td>0.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>2.3</td>
<td>Low</td>
<td>1.2</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>3.3</td>
<td>Medium</td>
<td>2.9</td>
<td>High</td>
<td>3.0</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td>1.5</td>
<td>Low</td>
<td>1.4</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6</td>
<td>2.4</td>
<td>Low</td>
<td>1.6</td>
<td>Low</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T7</td>
<td>2.8</td>
<td>Medium</td>
<td>1.7</td>
<td>Low</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T8</td>
<td>2.2</td>
<td>Low</td>
<td>1.7</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T9</td>
<td>3.9</td>
<td>High</td>
<td>2.5</td>
<td>Medium</td>
<td>3.0</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T10</td>
<td>3.2</td>
<td>Medium</td>
<td>2.0</td>
<td>Medium</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T11</td>
<td>4.4</td>
<td>High</td>
<td>2.5</td>
<td>Medium</td>
<td>3.5</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T12</td>
<td>2.9</td>
<td>Medium</td>
<td>2.1</td>
<td>Medium</td>
<td>2.0</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T13</td>
<td>2</td>
<td>Low</td>
<td>1.1</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T14</td>
<td>1.9</td>
<td>Low</td>
<td>1.3</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T15</td>
<td>2.4</td>
<td>Low</td>
<td>1.2</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T16</td>
<td>3.1</td>
<td>Medium</td>
<td>3.4</td>
<td>High</td>
<td>4.0</td>
<td>High</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.12 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO RESEARCH INTENSITY

All three calculations representing the central tendency with regard to the number of steps per research unit are combined below to provide the basis for allowing this aspect of the participants’ research behaviour to be assigned a Generous, Moderate or Frugal label. Data shows that half of the sample was characterised as Frugal and the rest of the sample was divided between the Moderate and Generous categories, with four participants in each group (see Table 7.13 below).

<table>
<thead>
<tr>
<th></th>
<th>1 (mean)</th>
<th></th>
<th></th>
<th>2 (weighted mean)</th>
<th></th>
<th></th>
<th>3 (median)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of steps per RU</td>
<td>Classification</td>
<td>Number of steps per RU</td>
<td>Classification</td>
<td>Number of steps per RU</td>
<td>Classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>2.9</td>
<td>Medium</td>
<td>1.8</td>
<td>Low</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>0.0</td>
<td>n/a</td>
<td>0.0</td>
<td>n/a</td>
<td>0.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>2.3</td>
<td>Low</td>
<td>1.2</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>3.3</td>
<td>Medium</td>
<td>2.9</td>
<td>High</td>
<td>3.0</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td>1.5</td>
<td>Low</td>
<td>1.4</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T6</td>
<td>2.4</td>
<td>Low</td>
<td>1.6</td>
<td>Low</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T7</td>
<td>2.8</td>
<td>Medium</td>
<td>1.7</td>
<td>Low</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T8</td>
<td>2.2</td>
<td>Low</td>
<td>1.7</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T9</td>
<td>3.9</td>
<td>High</td>
<td>2.5</td>
<td>Medium</td>
<td>3.0</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T10</td>
<td>3.2</td>
<td>Medium</td>
<td>2.0</td>
<td>Medium</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T11</td>
<td>4.4</td>
<td>High</td>
<td>2.5</td>
<td>Medium</td>
<td>3.5</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T12</td>
<td>2.9</td>
<td>Medium</td>
<td>2.1</td>
<td>Medium</td>
<td>2.0</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T13</td>
<td>2</td>
<td>Low</td>
<td>1.1</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T14</td>
<td>1.9</td>
<td>Low</td>
<td>1.3</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T15</td>
<td>2.4</td>
<td>Low</td>
<td>1.2</td>
<td>Low</td>
<td>1.0</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T16</td>
<td>3.1</td>
<td>Medium</td>
<td>3.4</td>
<td>High</td>
<td>4.0</td>
<td>High</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.13 RESEARCH INTENSITY CLASSIFICATION

Research intensity was also calculated by Nord (2009:205) at 1.8 steps per research unit (usage actions in one usage context in Nord’s terminology). The study was carried out using a sample of thirteen professional translators in “authentic translation situations [...] using aids normally available to them” (Nord, 2009:203). Taking into account that Nord’s study was carried out in 1997 and was, as far as can be ascertained, predominantly using paper resources, it is interesting to see that the research intensity between Nord’s study and the present study differs by one step (1.8 vs. 2.8 in both studies accordingly).
7.1.1.4. Research diversity

Research diversity is a category that provides insight into the range of resources used by individual translators. It does not make the distinction between the termino-lexicographic and text-based resources discussed in the previous chapter but considers the overall assortment of resources translators have in their tool box rather than their type or nature. The various aspects related to research diversity were discussed in Chapter 5 from the common tendencies perspective. Here, the research diversity category considers individual differences and distinguishes those translators who use few resources such as a dictionary, Google and Wikipedia, from those who use a much broader variety of resources, for example T15 who accessed twelve types of resources and fourteen specific resources in the translation task. Three types of calculations are considered in research diversity: the variety of types of resources accessed per task, the variety of specific resources consulted per task, and the number of specific resources used per research unit.

The variety of types of resources in the task was measured by the number of types of resources accessed by the participants during the translation task. Apart from T2 who did not use any resources, at the lower end of the spectrum there were six translators (T5, T6, T7, T8, T12 and T13) who tend to use between four and five resources, mostly dictionaries, search engines and parallel texts. On the high end of the spectrum, with as many as twelve different types of resources, was T15, who used a monolingual dictionary, a bilingual dictionary, a dictionary of synonyms, a dictionary of idioms, a glossary, a thesaurus, a discussion forum, a search engine, a knowledge-based resource, a concordancer, parallel texts and an online document. The participants were classified as having a Low, Medium or High number of types, depending on how many types they used in the task. As illustrated in Table 7.14 below, seven translators were classified as having a Medium number of types of resources in their tool box, five translators were classified as Low and three as High.

<table>
<thead>
<tr>
<th>Research diversity (the number of types of resources used in the task)</th>
<th>Number of types of resources</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>10</td>
<td>High</td>
</tr>
<tr>
<td>T2</td>
<td>0</td>
<td>Low</td>
</tr>
<tr>
<td>T3</td>
<td>7</td>
<td>Medium</td>
</tr>
<tr>
<td>T4</td>
<td>6</td>
<td>Medium</td>
</tr>
<tr>
<td>T5</td>
<td>4</td>
<td>Low</td>
</tr>
<tr>
<td>T6</td>
<td>4</td>
<td>Low</td>
</tr>
<tr>
<td>T7</td>
<td>5</td>
<td>Medium</td>
</tr>
<tr>
<td>T8</td>
<td>4</td>
<td>Low</td>
</tr>
<tr>
<td>T9</td>
<td>7</td>
<td>Medium</td>
</tr>
<tr>
<td>T10</td>
<td>7</td>
<td>Medium</td>
</tr>
<tr>
<td>T11</td>
<td>9</td>
<td>High</td>
</tr>
<tr>
<td>T12</td>
<td>5</td>
<td>Medium</td>
</tr>
<tr>
<td>T13</td>
<td>4</td>
<td>Low</td>
</tr>
<tr>
<td>T14</td>
<td>7</td>
<td>Medium</td>
</tr>
<tr>
<td>T15</td>
<td>12</td>
<td>High</td>
</tr>
</tbody>
</table>
The variety of specific resources was already discussed in 7.1.1.2. as this calculation was also used in establishing resource volume. It is measured by the number of specific resources accessed by the participants during the translation task. Table 7.7 in section 7.1.1.2. illustrates the classification of translators into the Low, Medium or High variety of specific resources use, depending on the number of specific resources they accessed during the translation task.

The number of specific resources per RU also contributes to establishing research diversity as it shows the variety of specific resources used in each research episode rather than in the whole translation task. In order to classify translators into three groups depending on whether they used a Low, Medium or High number of specific resources per RU, it was necessary to establish a central tendency, that is, average values indicating how many specific resources, on average, were used by individual translators in the translation task. Only the weighted mean was taken into account in this calculation as ranges were too narrow for the mean and median to provide a meaningful contribution to the calculation. The range was established to be 1.3, with the lowest value in the range being 1.4 (T5 and T8) and the highest 2.7 (T11) specific resources per RU. T2 was excluded from the calculations but assigned the Low category in keeping with his observed behaviour. The overall tendency seems to be towards a Low/Medium number of specific resources per RU as seven participants were classified as Low and Medium and two as High.

<table>
<thead>
<tr>
<th>Research diversity (the number of specific resources per RU)</th>
<th>Number of specific resources per RU</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>1.9</td>
<td>Medium</td>
</tr>
<tr>
<td>T2</td>
<td>0.0</td>
<td>Low</td>
</tr>
<tr>
<td>T3</td>
<td>1.5</td>
<td>Low</td>
</tr>
<tr>
<td>T4</td>
<td>2.1</td>
<td>Medium</td>
</tr>
<tr>
<td>T5</td>
<td>1.4</td>
<td>Low</td>
</tr>
<tr>
<td>T6</td>
<td>1.6</td>
<td>Low</td>
</tr>
<tr>
<td>T7</td>
<td>1.9</td>
<td>Medium</td>
</tr>
<tr>
<td>T8</td>
<td>1.4</td>
<td>Low</td>
</tr>
<tr>
<td>T9</td>
<td>2.1</td>
<td>Medium</td>
</tr>
<tr>
<td>T10</td>
<td>2.2</td>
<td>Medium</td>
</tr>
<tr>
<td>T11</td>
<td>2.7</td>
<td>High</td>
</tr>
<tr>
<td>T12</td>
<td>1.9</td>
<td>Medium</td>
</tr>
<tr>
<td>T13</td>
<td>1.6</td>
<td>Low</td>
</tr>
<tr>
<td>T14</td>
<td>1.7</td>
<td>Low</td>
</tr>
<tr>
<td>T15</td>
<td>1.9</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Table 7.14 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NUMBER OF TYPES OF RESOURCES USED IN THE TRANSLATION TASK
Table 7.15 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE NUMBER OF SPECIFIC RESOURCES USED PER RESEARCH UNIT

Table 7.15 illustrates how all three aspects of research diversity are combined to classify participants into three categories: Frugal, Moderate or Generous depending on whether they were classified as having a Low, Medium or High number of resources in the three aspects discussed so far in this section. Nearly half of the translators proved to be Frugal, 43% Moderate and just over 12% Generous. The data analysis, therefore, reveals a range of diversity amongst the translators in this sample, but the overall tendency is towards the Low/Medium diversity.

Table 7.16 CLASSIFICATION OF TRANSLATORS ACCORDING TO RESEARCH DIVERSITY

7.1.1.5. Research time

The amount of time translators are willing to devote to research is an important aspect of their research style. Whilst translators with explorative tendencies will devote a generous amount of time to research, others, with minimalistic inclinations, will keep research time to an absolute minimum. It can be argued that under time pressure these differences might be diminished. However, it is unlikely that they can be completely removed. An example of such variations, albeit from a slightly different perspective, can be seen in the recent discussion on the translators’ forum ProZ (2015a:online), where translators gave their opinions about how many words they could translate in eight hours. It is interesting to see that while some said that they had “no idea how people can translate more than 1,500 words a day”, for others, 5,000 words a day was not a problem and they claimed to “do it all the time... even more, including checking afterwards and delivering final work of the highest quality” (ibid.). The translation task in the current study was performed by the participants without a time constraint to allow their natural predispositions to be revealed; however, one translator (T12) self-imposed a time limit in order to perform the task during the lunch break.

In the present research, two types of calculations were included in the research time category. The first type refers to the actual time spent on research and the second takes into account the proportion of time spent on research in relation to the time spent on translation (i.e. interaction with texts). The actual time was calculated by adding up the duration of all research episodes; the proportion of time spent on research was calculated as a percentage in relation to the time spent on interaction with the texts, also calculated as a percentage. Both calculations excluded time spent on commenting, where no interactions with texts or resources were taking place.
In Table 7.17 below, the actual research time is presented for all individual translators, together with the classification into High, Medium or Low research time based on the combined duration of all their research episodes. The range of the actual research time was surprisingly wide, with the lowest research time being 0’0’’ for T2, who did not conduct any research, and the highest 1h2’58’’ for T1\textsuperscript{214}. One of these participants, T1, was considered to be an outlier and, in line with the explanation given in section 3.4., the range was trimmed to exclude T1’s research time. Therefore, the range on the basis of which the translators were divided into High, Medium or Low, spans from 0’0’’ (T2) to 38’14’’ (T15). Although T1’s research timing was ignored in the range, her behaviour was still classified as High. All in all, the tendency seems to be towards the Low/Medium research time. Only three translators, whose research lasted for over 30’, were classified as High.

<table>
<thead>
<tr>
<th>Research time (actual)</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 01:02:58</td>
<td>High</td>
</tr>
<tr>
<td>T2 00:00:00</td>
<td>Low</td>
</tr>
<tr>
<td>T3 00:24:51</td>
<td>Medium</td>
</tr>
<tr>
<td>T4 00:14:05</td>
<td>Medium</td>
</tr>
<tr>
<td>T5 00:06:40</td>
<td>Low</td>
</tr>
<tr>
<td>T6 00:22:46</td>
<td>Medium</td>
</tr>
<tr>
<td>T7 00:16:51</td>
<td>Medium</td>
</tr>
<tr>
<td>T8 00:11:36</td>
<td>Low</td>
</tr>
<tr>
<td>T9 00:16:49</td>
<td>Medium</td>
</tr>
<tr>
<td>T10 00:24:11</td>
<td>High</td>
</tr>
<tr>
<td>T11 00:30:39</td>
<td>High</td>
</tr>
<tr>
<td>T12 00:08:53</td>
<td>Low</td>
</tr>
<tr>
<td>T13 00:12:39</td>
<td>Low</td>
</tr>
<tr>
<td>T14 00:13:11</td>
<td>Low</td>
</tr>
<tr>
<td>T15 00:38:14</td>
<td>High</td>
</tr>
<tr>
<td>T16 00:06:12</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 7.17 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO TIME SPENT ON RESEARCH (MEASURED IN REAL TIME)

The second calculation showing the research time measured in percentages in relation to the translation time is shown below in Table 7.18. As in the previous calculation relating to the actual research time, the data relating to outlier behaviour was trimmed from the dataset. Here, only T2 was excluded from the range as his 0% research time in relation to 100% translation time was removed from the mean further than the required 2.5 z-scores. Nevertheless, T2 was retained in the sample and was classified in the Low category based on the low end of the range his behaviour belonged to. The range was therefore 33%, with

\textsuperscript{214} T1’s Internet connection was slow, which was a contributing factor to the overall length of her translation task.
the lowest observed value of 23% (T5 & T8) and the highest 56% (T1). Individual translators were classified as having a Low, Medium or High percentage of research time depending on where they fell within the range (see Table 7.18).

As mentioned in section 5.2.1., translators participating in the translation task of the Main Study spent on average 33% of their time on translation-related research, with translators specialising in the economic/business/financial domains averaging 30% and those who did not, averaging 36%. This calculation revealed that when it comes to the time devoted to research activities in relation to translation activities, over half of the sample (56%) was classified as Low, compared to the previous calculation (actual research time), in which 44% of the sample was classified as having Low research time when measured in time units. Table 7.18 below shows the proportion of research time for each translator and how they were classified.

<table>
<thead>
<tr>
<th>Research time (%)</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 56%</td>
<td>High</td>
</tr>
<tr>
<td>T2 0%</td>
<td>Low</td>
</tr>
<tr>
<td>T3 36%</td>
<td>Medium</td>
</tr>
<tr>
<td>T4 33%</td>
<td>Low</td>
</tr>
<tr>
<td>T5 23%</td>
<td>Low</td>
</tr>
<tr>
<td>T6 36%</td>
<td>Medium</td>
</tr>
<tr>
<td>T7 41%</td>
<td>Medium</td>
</tr>
<tr>
<td>T8 23%</td>
<td>Low</td>
</tr>
<tr>
<td>T9 43%</td>
<td>Medium</td>
</tr>
<tr>
<td>T10 47%</td>
<td>High</td>
</tr>
<tr>
<td>T11 32%</td>
<td>Low</td>
</tr>
<tr>
<td>T12 30%</td>
<td>Low</td>
</tr>
<tr>
<td>T13 30%</td>
<td>Low</td>
</tr>
<tr>
<td>T14 28%</td>
<td>Low</td>
</tr>
<tr>
<td>T15 49%</td>
<td>High</td>
</tr>
<tr>
<td>T16 27%</td>
<td>Low</td>
</tr>
</tbody>
</table>

Range: 33% (23% - 56%)
23%-34% Low
35%-46% Medium
47%-56% High

Table 7.18 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO TIME SPENT ON RESEARCH (MEASURED IN PROPORTION TO TRANSLATION TIME)

Figure 7.1 below brings both calculations relating to research time together to show their combined effect on the research time classification. The Moderate and Rapid labels have been given to seven translators respectively and only two participants (T1 and T15) were labelled as Relaxed. As previously, in cases where different classifications were assigned to the two calculations (e.g. T7 and T9), the calculation with the lowest standard deviation within its range was accepted as dominant.
7.1.1.6. Research pace

Examining research pace provides insight into the tempo of the translator’s progression through research steps in a research episode. While some translators take their time engaging with individual resources as their research progresses, other translators’ research is characterised by fleeting encounters with resources in a given research episode. Research pace provides a different perspective to research time as it is tied to the number of steps performed during the research episodes and the duration of these steps and episodes. By considering research pace, we do not look at how long participants took to satisfy all their research needs overall; we are interested in how quickly each of these needs was satisfied. Two calculations were performed to establish translators’ research pace: the average duration of research step and the average duration of research episode.

It was established in section 5.3.1. that an average research step lasted 20 seconds, with a range of 26, spanning between 12 and 38 seconds. This calculation excluded T2 due to the different, group perspective being sought. Here, T2 is also excluded from the calculations in line with the adopted treatment of outliers (see section 3.4.). All participants were classified into three categories as having Short, Medium or Long average research steps, depending on the average length of their observed research step. Most of the participants’ research (over 69% of the sample) was characterised by research steps of Short duration, three translators were classified in the Medium and two in the Long category (see Table 7.19).

<table>
<thead>
<tr>
<th>Translator</th>
<th>Average duration of research step (actual)</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>00:00:38</td>
<td>Long</td>
</tr>
<tr>
<td>T2</td>
<td>00:00:00</td>
<td>Short</td>
</tr>
<tr>
<td>T3</td>
<td>00:00:16</td>
<td>Short</td>
</tr>
<tr>
<td>T4</td>
<td>00:00:21</td>
<td>Medium</td>
</tr>
<tr>
<td>T5</td>
<td>00:00:29</td>
<td>Medium</td>
</tr>
<tr>
<td>T6</td>
<td>00:00:20</td>
<td>Short</td>
</tr>
<tr>
<td>T7</td>
<td>00:00:34</td>
<td>Long</td>
</tr>
<tr>
<td>T8</td>
<td>00:00:19</td>
<td>Short</td>
</tr>
<tr>
<td>T9</td>
<td>00:00:12</td>
<td>Short</td>
</tr>
<tr>
<td>T10</td>
<td>00:00:17</td>
<td>Short</td>
</tr>
<tr>
<td>T11</td>
<td>00:00:14</td>
<td>Short</td>
</tr>
<tr>
<td>T12</td>
<td>00:00:17</td>
<td>Short</td>
</tr>
<tr>
<td>T13</td>
<td>00:00:18</td>
<td>Short</td>
</tr>
</tbody>
</table>
The second aspect of the research pace category examined here is the average duration of research episodes, which was calculated by dividing the translator’s total research time by the number of their research episodes. Here, T2 was included in the calculations as the distance between the mean (56”) and T2’s zero research episode time was smaller than the required 2.5 T-scores (01’7”). Therefore, the range is calculated at 1’45”.

The participants were classified into three categories corresponding to Short, Medium or Long average duration of research episodes. Most of the translators (over 62% of the sample) were characterised as having Medium duration of research episodes. Four translators were classified as having Long and only two as having Short duration of research episodes. This would mean that although the constituent steps of research episodes are mostly short, the research episodes themselves are mostly medium in length.

| T14 | 00:00:20 | Short |
| T15 | 00:00:24 | Medium |
| T16 | 00:00:13 | Short |

Range: 26’’ (0’12” – 0’38’’)

| 00:00:00 - 00:00:20 | Short |
| 00:00:21 - 00:00:29 | Medium |
| 00:00:30 - 00:00:38 | Long |

Table 7.19 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO THE AVERAGE DURATION OF RESEARCH STEPS

<table>
<thead>
<tr>
<th>Average duration of research episode (actual)</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average duration of research episode hrs:min:sec</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>00:01:45</td>
</tr>
<tr>
<td>T2</td>
<td>00:00:00</td>
</tr>
<tr>
<td>T3</td>
<td>00:00:32</td>
</tr>
<tr>
<td>T4</td>
<td>00:01:17</td>
</tr>
<tr>
<td>T5</td>
<td>00:00:50</td>
</tr>
<tr>
<td>T6</td>
<td>00:01:45</td>
</tr>
<tr>
<td>T7</td>
<td>00:01:32</td>
</tr>
<tr>
<td>T8</td>
<td>00:00:39</td>
</tr>
<tr>
<td>T9</td>
<td>00:00:44</td>
</tr>
<tr>
<td>T10</td>
<td>00:00:58</td>
</tr>
<tr>
<td>T11</td>
<td>00:00:54</td>
</tr>
<tr>
<td>T12</td>
<td>00:00:53</td>
</tr>
<tr>
<td>T13</td>
<td>00:00:36</td>
</tr>
<tr>
<td>T14</td>
<td>00:00:46</td>
</tr>
<tr>
<td>T15</td>
<td>00:01:00</td>
</tr>
<tr>
<td>T16</td>
<td>00:00:41</td>
</tr>
</tbody>
</table>
The two aspects of research pace are combined below in Table 7.21 to show how they affect the classification of the participants as having a Relaxed, Moderate or Rapid pace of research. As before, the category with the lower standard deviation within its range is considered more dominant. Most translators appear to have Moderate research pace (over 62% of the sample), followed by the Relaxed type (25% of the sample). Only two translators had Rapid research pace, including T2, who did not conduct any research. This would point to the fact that most translators participating in this study had a tendency towards a Moderate pace of research, which means that they progressed through research in a steady manner.

7.1.1.7. Primary categories summary

As mentioned earlier, the primary categories are a fundamental part of the TTRS. They are clustered into volume-related and time-related features. This clustering was performed in order to place translators on a grid which I have called the TTRS grid, with one axis representing the volume of the research undertaken, and the other representing the time-related features of this research. Table 7.22 and Table 7.23 below show this clustering, with Table 7.22 representing the Volume-related axis and Table 7.23 representing the Time-related axis.
The TTRS labels in both axes were assigned the following weight scores in order to create ranges for both axes which would allow translators to be placed on the TTRS grid:

- Frugal and Rapid - a score of 2
- Moderate - a score of 4
- Generous and Relaxed - a score of 6.

The final weight scores for both axes shown in Table 7.24 correspond to the ranges on the TTRS grid.

The scores were calculated for each translator for each of the axis giving a final weight sum, which created a range of scores representing the range of characteristic behaviours (see Table 7.24). For the Volume axis, the range was 8 to 24 (the lowest possible score would be 2 x 4 = 8 and the highest would be 6 x 4 = 24 as there are four categories related to the Volume axis and the lowest score is 2 for Frugal and 6 for Generous). For the Time axis, the range was 4 to 12 (the lowest possible score would be 2 x 2 = 4 and the highest would be 6 x 2 = 12 as there are two categories related to the Time axis and the lowest score is 2 for Rapid and 6 for Relaxed).

### Table 7.24 Score Calculation for the Volume and Time Axis of the TTRS

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ru volume</td>
<td>Frugal</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource volume</td>
<td>Frugal</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Diversity</td>
<td>Frugal</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Intensity</td>
<td>Frugal</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frugal weight score</td>
<td>Frugal</td>
<td>0</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Ru volume</td>
<td>Moderate</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource volume</td>
<td>Moderate</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Diversity</td>
<td>Moderate</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Intensity</td>
<td>Moderate</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate weight score</td>
<td>Moderate</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>12</td>
<td>16</td>
<td>0</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Ru volume</td>
<td>Generous</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource volume</td>
<td>Generous</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Diversity</td>
<td>Generous</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Intensity</td>
<td>Generous</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generous weight sum</td>
<td>Generous</td>
<td>18</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Final weight score</td>
<td>Generous</td>
<td>22</td>
<td>8</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>8</td>
<td>8</td>
<td>12</td>
<td>18</td>
<td>18</td>
<td>16</td>
<td>24</td>
<td>12</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>LABEL</td>
<td>Generous</td>
<td>G</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>G</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research time</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research pace</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Final weight score</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LABEL</td>
<td>Relaxed</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research time</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research pace</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Final weight score</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LABEL</td>
<td>Relaxed</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RX</td>
<td>RD</td>
<td>RD</td>
<td>RD</td>
</tr>
</tbody>
</table>

Table 7.24 Score Calculation for the Volume and Time Axis of the TTRS
Based on the score table above, the individual translators were placed on the TTRS grid according to their final weight scores. The grid contains five areas corresponding to the five different ways of conducting research which emerged from the data (prolific, explorative, moderate, economical and understated). The translators who fitted into one of the delineated areas were associated with the corresponding research behaviour. For example, T1’s final Volume score was 22 and her final Time score was 12. She was therefore placed on the TTRS grid in the bottom right corner (see Figure 7.2). Five translators (T3, T4, T11, T15 & T16) were plotted on the border positions and their positions then established based on the secondary categories.

Figure 7.2 THE TTRS GRID SHOWING THE POSITIONING OF INDIVIDUAL TRANSLATORS ACCORDING TO THEIR WEIGHT SCORES FOR PRIMARY CATEGORIES

In the following section, we will discuss the secondary categories which will contribute to these initial research profiles by either strengthening their position or shifting it towards one of the neighbouring ones.

7.1.2. Secondary categories (attributes)

In contrast to the primary categories, the fourteen secondary categories (listed in section 7.1.) are of a qualitative nature and are mostly based on repetitive occurrences of certain observations from the screen recordings (two of the secondary categories, information retention and attitude towards technology, are based on self-declared data from the profile questionnaires). Unlike many of the primary categories, the secondary categories represent only one aspect of translators’ online research behaviour. Because they are characteristic of some participants, but not the others, they are regarded as attributes rather than categories per se and are assigned to a particular translator’s research profile in a graded fashion, depending on the strength of their presence in that translator’s research. If a particular behaviour was

---

215 The two categories based on self-declared data (information retention and attitude towards technology) do not represent research behaviour as such, but are related to it.
observed, then the attribute was assigned to the participant’s research profile as an additional feature of their research patterns. This was done by assigning weights to the observed behaviour phenomena. As shown in Table 7.25, these weight scores created ranges, which were then split into three ‘gradients of presence’ - weakly present (WP), moderately present (MP) and strongly present (SP). The weak presence received a weight of 1, the moderate presence a weight of 2 and the strong presence a weight of 3.

<table>
<thead>
<tr>
<th>Behaviour attribute</th>
<th>Absent</th>
<th>Weakly Present</th>
<th>Moderately Present</th>
<th>Strongly Present</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight assigned</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 7.25 WEIGHT SYSTEM FOR ASSIGNING SECONDARY CATEGORIES (ATTRIBUTES) TO TRANSLATORS’ PROFILES

7.1.2.1. Deep searches

Research depth constitutes an important dimension of the research profiles of professional translators as it indicates their level of engagement with online resources. The depth of research is one of the recurring features of information behaviour studies, (e.g. Heinström, 2005, 2006b; Enríquez Raído, 2011). It was originally envisaged to be a primary category; however, because it did not fit into either of the axes (volume or time related) it is considered here as the most important secondary category. Research depth is tied to a research episode. While some translators’ research episodes are characterised by brief, cursory consultations, others engage with resources more deeply. In this study, deep searches are those research episodes in which the translator is observed to engage with online resources for a longer period of time or performs a chain of online interactions, often modifying queries and changing the type of resource in pursuit of an acceptable solution to the problem.

Two translators (T6 and T11) were observed to deeply engage with their resources in over a third of their research episodes and four carried out deep research in 20-30% of their research episodes. Translators who had less than 18% of deep searches were considered to have a weak presence of deep searches in the translation task (see Table 7.26 below).

<table>
<thead>
<tr>
<th>No of deep searches</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of deep searches in relation to all research episodes</td>
<td>29%</td>
<td>0%</td>
<td>9%</td>
<td>18%</td>
<td>13%</td>
<td>38%</td>
<td>27%</td>
<td>11%</td>
<td>9%</td>
<td>24%</td>
<td>34%</td>
<td>20%</td>
<td>10%</td>
<td>11%</td>
<td>14%</td>
<td>0%</td>
</tr>
<tr>
<td>Attribute grading label</td>
<td>SP</td>
<td>MP</td>
<td>WP</td>
<td>WP</td>
<td>SP</td>
<td>WP</td>
<td>WP</td>
<td>SP</td>
<td>SP</td>
<td>WP</td>
<td>WP</td>
<td>WP</td>
<td>WP</td>
<td>WP</td>
<td>WP</td>
<td>A</td>
</tr>
<tr>
<td>Weight assigned</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Range</td>
<td>29%</td>
<td>9%</td>
<td>38%</td>
<td>0%</td>
<td>1%</td>
<td>2%</td>
<td>1%</td>
<td>3%</td>
<td>2%</td>
<td>3%</td>
<td>2%</td>
<td>1%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.26 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO DEEP SEARCHES

7.1.2.2. Shallow searches

In contrast to deep searches, shallow searches are often quick consultations in termino-lexical resources or a rapid Google search with snippet viewing. Four translators were found to have a high percentage of shallow searches, with research episodes characterised by one-step or fleeting consultations constituting 60-89% of their research episodes. Two translators’ research episodes were characterised as having a
A moderate number of shallow searches and eight were considered to have a weak presence of shallow research episodes (see Table 7.27 below).

It has to be said that translators normally have a mix of deep, moderate or shallow research episodes, depending on the difficulty of a particular research unit or on how quickly they manage to satisfy their research need. However, data presented in this study shows that while some participants did not have any deep episodes (T16), some did not have any shallow ones (T6). Therefore, the depth of research which a translator is prepared to undertake with regard to research needs is also something that has to be considered as part of one’s research style.

Table 7.27 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO SHALLOW SEARCHES

When comparing these results with findings from general information behaviour, i.e. not specific to translator behaviour, it seems that translators’ search behaviour does not follow the same patterns. For example, Nicholas et al. (2006) measured the level of people’s engagement with websites using a website penetration metric and found it to be mostly shallow, with visits to websites being “brief, to the point and possibly cursory” (Nicholas et al., 2006:210). Similarly, the ‘Google generation’ study (Rowlands et al., 2008) has found general online behaviour to be rather shallow and horizontal, with lots of skimming activities. Translators in the present study, however, exhibited patterns of a more engaged nature as over a third of the sample displayed a medium or strong presence of deep searches. However, as mentioned before, in the present study the depth of research was assessed as a combined effect of the penetration of a particular website and the length of the research episodes, and therefore, cannot be directly compared to the findings from general studies. In translation studies, Enríquez Raído (2011:362) considered research depth, but her analysis of the depth included additional aspects of online behaviour such as “the number and type of internal site queries conducted, internal and external pages accessed, [...] overall time spent consuming website content [...] [and] number and type of search engine query modifications”. However, in Enríquez Raído’s study all the above-mentioned aspects were not systematically compared as a combination of factors contributing to the depth of search. Rather, those individual aspects were qualitatively described with the support of descriptive statistics for each of the six participants, concluding that the searches of novice translators were of a mostly ‘shallow’ or ‘checking’ nature (Enríquez Raído, 2011:360-405). Therefore, the findings of Enríquez Raído cannot be directly compared to the findings of the present study either, not only because of the different focus of her study but also because a different method of establishing depth was used. Nevertheless, it is worth pointing out that the differences in depth of research between the novice and experienced translators observed by Enríquez Raído can also be seen amongst the group of professional translators studied here, as the depth of searches varied considerably between the participants in this study. This raises a question of whether the depth of searches can be solely attributed to experience or whether it could also depend on research style.
7.1.2.3. Meandering research path

A research path can be described as a trajectory of the sequence of steps within a research episode and, in this study, refers to the shape of this trajectory. Research paths are an important component of the Typology of Translator Research Styles as they indicate translators’ predispositions towards undertaking research in a certain fashion that can be linked to their level of exploratory tendencies. Most translators proceed with their research in a straight path, i.e. they consult one resource after another, without going back to previously consulted resources or branching out. However, a small number of participants were observed to diverge frequently from the straight path by following the internal links and hyperlinks of a resource, usually in order to obtain additional information. They were also observed to often engage with resources for longer periods of time, which suggests reading rather than skimming, and taking a deeper interest in the subject matter. This kind of behaviour was labelled as *meandering research path* and is associated with explorative tendencies. This kind of behaviour often results in indirect research (see section 5.3.2.).

Search paths could only be observed in the longer research episodes containing four or more steps. Therefore, only 72 straight and 11 meandering paths were observed from the whole sample out of the total of 321 research episodes. Since all translators were observed to follow straight research paths, they were not considered to be a ‘distinguishable’ feature. Therefore, only the meandering path trajectories are considered as an attribute. Table 7.28 below illustrates the observed instances of meandering research paths in absolute and relative frequencies (percentages). Table 7.28 also shows how attribute grading labels (Absent, Weakly present, Moderately present and Strongly present) were attached to those profiles which exhibited meandering behaviour, according to the grading system.

<table>
<thead>
<tr>
<th>Translator</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observed instances</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>% of all observed longer paths</td>
<td>12%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>10%</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>0%</td>
<td>55%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>20%</td>
</tr>
<tr>
<td>Attribute grading label</td>
<td>WP</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>WP</td>
<td>A</td>
<td>A</td>
<td>WP</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>WP</td>
<td></td>
</tr>
<tr>
<td>Weight attached</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.28 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO MEANDERING RESEARCH PATHS

The most meandering translator was T11, who frequently followed internal links and hyperlinks, her searches often branching out. Four other translators’ research paths were characterised as meandering to a lesser degree. Meandering paths also included, for example, researching the author of a publication or other related parallel texts and reading more information than required for the translation purpose, usually out of interest etc. For example, after having consulted an article for one minute and 15 seconds, T6 commented in the audio commentary that she was aware that reading the rest of the article was a loss of

---

216 The only exceptions are searches initiated with search engines, where clicking on a result is not considered branching out; however, any subsequent clicking on links from within this resource is.

217 Apart from T5, whose research episodes were too short to observe research paths.
time for translation but she was starting to find it interesting. She then went on reading it for another 30 seconds.

7.1.2.4. Parallel research sequence

While the research path is concerned with the shape of the research trajectory, research sequence refers to the manner of viewing of the resources in a research episode. Most translators usually follow what I have called a sequential pattern, whereby resources are consulted one after another, i.e. when an answer is not found in the first consulted resource, a translator moves on to the second, then third etc. However, some translators were observed to perform their search in what I have called a parallel research sequence, whereby a translator consults more than one resource at the same time, often by means of right-clicking on search engine results and opening two, three or more results in a new window for consultation. While sequential research indicates following a ‘one at a time’ principle in which the translator hopes to find the answer in the first one consulted, in parallel research, opening several resources one after the other before consulting them suggests a presupposition that more than one resource will be needed to satisfy a research need. This type of parallel research could be indicative of a need to economise on time but could also indicate explorative tendencies.

Similarly to research paths, research sequence could only be observed in the longer research episodes containing more than four research steps. Therefore, only 28 parallel and 68 sequential patterns were identified. All translators but T5 were observed to access online resources in a sequential manner. Therefore, this was not considered a ‘distinguishable behaviour’ and only parallel research sequences are considered here. Table 7.29 below shows only the instances of parallel research sequences measured in absolute and relative frequencies (percentages). It also shows how attribute grading labels (Absent, Weakly present, Moderately present and Strongly present) were attached to those profiles which exhibited parallel sequence behaviour, according to the grading system.

![Table 7.29 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO RESEARCH SEQUENCE](image)

T11 was the most prolific parallel searcher and performed 21 searches of this type during the translation task, which accounted for 88% of all her research episodes containing four steps or more. Four other translators’ research was characterised as displaying a parallel sequence pattern to a lesser degree.
7.1.2.5. Planning

Planning, in this study, refers to any type of consideration on the part of the translator of what research activities will be required for the task before proceeding with the translation. This might include reading or skimming the text and highlighting words that will require research (as T4 and T9 have done) or carrying out an initial assessment of the main concepts or ideas that will need researching (as T9 and T10 have done). Planning could also be associated with concepts introduced by Carl et al., (2011, np) referring to a translator’s working style in the initial, i.e. orientation phase. These are: “systematic initial orientation, where the translator systematically reads through the ST before translation [or] skimming, where translator skims the ST rapidly before translation”. Dragsted & Carl observed three of their participants to have a much longer orientation phase than others and thus engaging with the text more intensely at the beginning (2013:141). However, since no use of external resources was allowed in their study, it is difficult to know whether these participants would have engaged in planning research activities at this stage. This type of behaviour can often be associated with those translators in the present study who prefer front-loaded clustering of research (see section 5.3.2.).

In the profile questionnaire, translators were asked to declare what type of initial orientation they normally carry out in their everyday work. They could choose from: systematic orientation (reading the text), skimming (rapid skim), quick planning (reading a couple of first sentences) or head start (starting to translate right away). Table 7.30 below compares the self-declared initial orientation with that observed in the screen recordings\(^{218}\). Most of the translators who were observed to perform systematic orientation said they normally skim the text before translating. Those translators who were observed to carry out a planning activity in the form of reading the text and highlighting words requiring research received a weight of 3 and those who read the text but did not highlight any words received a weight of 2\(^{219}\).

| Self-declared                                      | T1   | T2   | T3   | T4   | T5   | T6   | T7   | T8   | T9   | T10  | T11  | T12  | T13  | T14  | T15  | T16  |
|---------------------------------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Head start                                        | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 3    | 3    | 0    | 0    | 0    | 0    | 0    |     |
| Skimming                                          | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    |     |
| Quick planning                                    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    |     |
| Systematic orientation                           |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Orientation with highlighting                     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Classifications of individual translators         |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Weight attached                                    | 0    | 0    | 2    | 3    | 0    | 0    | 0    | 0    | 0    | 3    | 3    | 0    | 0    | 0    | 0    | 0    |
| Attribute grading                                 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Head start/skimming/quick planning                | 0    | 0    | 2    | 3    | 0    | 0    | 0    | 0    | 0    | 3    | 3    | 0    | 0    | 0    | 0    | 0    |
| Systematic orientation without highlighting       | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 2    |
| Systematic orientation with highlighting          | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 3    |

Table 7.30 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO PLANNING ATTRIBUTE

Three translators were classified as strong planners and two as planning to a lesser degree by familiarising themselves with the text. It has to be said, however, that because of the nature of this study, it is not possible to confirm that participants had not read the text before setting out to carry out the task. Although subjects received detailed instructions, it was not explicitly stated that they should only open the text at the start of the task. Therefore, pre-consultations cannot be precluded.

---

\(^{218}\) The initial orientation of T8 is not known as it was not present in the screen recording. The attempt to establish this through the post-task questionnaire also failed as the answer was vague (i.e. the answer to the question “Did you consult the text before you started translating it? Did you read it, skim it, read the first few sentences or did you start translating without reading the text?” was: “Yes, I did”.

\(^{219}\) Since only two options were considered, two weights out of four (0, 1, 2 or 3) had to be assigned. Because planning is considered to be an important factor in translator research behaviour, it was decided that the two highest weights (2 &3) should be used.
7.1.2.6. Squirrelling

The term ‘squirrelling’ was borrowed from information behaviour literature where it originally referred to a strong consumer instinct demonstrated by storing away content in the form of downloads (Rowlands et al., 2008:295). In the context of this study, squirrelling refers to any activity related to saving information for future reference during the translation process, whether by creating a new entry in a termbase or adding a term to a personal glossary or a Word dictionary. Table 7.31 below illustrates to which individual translators squirrelling behaviour was attributed.

Table 7.31 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO SQUIRRELLING BEHAVIOUR ATTRIBUTE

Three participants (T1, T6 and T11) displayed this type of behaviour, with T6 showing the strongest inclination to save information in her Déjà Vu termbase. This is thought-provoking because T6 explicitly said in the audio commentary that it is unlikely she would translate such a text in future as she “doesn’t do this kind of translation’ and yet, she created eleven entries in her termbase during the translation task. This would indicate that squirrelling behaviour is symptomatic of research style as it has been demonstrated that it can occur irrespective of whether the information stored is thought to be useful for the future.

7.1.2.7. Snippet viewing

Snippet viewing refers to consulting the results of a search, i.e. the information contained in the page title, the description text appearing below links in search results and possibly the URL, usually in a search engine, without actually following any of the links. This behaviour is distinguished from standard browsing, which is associated with following links and hyperlinks, by the fact that no links are followed during the viewing. Table 7.32 below illustrates to which individual translators snippet viewing was attributed.

---

220 T6 referred to this action as ‘saving to memory’, which, in fact it is creating a new entry in the termbase of her CAT Tool, Déjà Vu.

221 Enríquez Raído (2011:293) uses the acronym SERPs (search engine results pages) to refer to the information displayed on Google’s search results pages. This information includes the title of the Web pages (i.e. the first line of any search result), the snippet (the short description of a Web page, which may include an excerpt from the page), and the URL or address of a result’s Web page.
A total of five translators were observed to perform this kind of search action, with T9 being the most prolific snippet viewer. T9’s snippet viewing behaviour is also related to her hybrid type of research strategy (see section 6.1.3.), i.e. the combination of top-down and bottom-up approaches, which is clearly intended to deliver the required information without the need for accessing any websites. Snippet viewing is associated with the quick and often shallow research patterns; however, it can also be used by the more explorative translators to widen their search by scanning pages of results, sometimes going beyond the third page of results.

The combination of the hybrid research strategy, advanced search and snippet viewing can be seen as a kind of ‘speed search’ and, as the content of linguistic resources becomes more embedded in the web searches, this type of search might be a solution for a super-fast and reliable retrieval of information for translation purposes in the future.

7.1.2.8. Using advanced queries

Using advanced queries in this study is understood as “using options offered by search engines to efficiently, effectively, and directly control the quality of search results. Direct control and implied conscious action by the user are critical to advanced searching” (Hock, 2008:14). Advanced searching implies advanced procedural knowledge of web-searching techniques, which is vital for the possession of the instrumental sub-competence as defined by PACTE (2009:208). In other words, as Hock suggests, it requires “an ‘advanced searcher,’ someone who brings to the search a particular combination of knowledge, skills, attitudes, and perspectives” (Hock, 2008:14) in order to formulate a query that reflects their research needs. However, it can be argued that the use of advanced queries is not just a reflection of the user’s web-searching expertise; it can also be associated with research styles.

Main Study participants were asked whether they use advanced search features such as Boolean logic truncation, wildcard, nested search or proximity search. All translators declared using at least one of these features. However, in the translation task 38% of the respondents did not use any advanced search features. This could partly be due to the fact that the text did not necessitate the use of such features for some of the participants, especially for T2 who did not carry out any research and T5, T12 and T16 who had few research needs. However, certain advanced search features were used by the participants in the translation task, including exact phrase (T3, T4, T6, T7, T9, T10, T15), language selection (T3, T7, T11), define function (T10), define + a filter word (T10, T6, T4), exact phrase + file type (T4), exact phrase + domain (T3). These functions were mostly used in search engine searches but were also applied within particular resources, e.g. ProZ glossaries. Table 7.33 below shows which translators used advanced queries
and what percentage of all their research needs were subjected to advanced search. Thus, the most prolific users of advanced search were T4 and T10, followed by T3 and T15. The remaining six translators used advanced search to a lesser degree.

<table>
<thead>
<tr>
<th>Advanced queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of observed advanced queries</td>
</tr>
<tr>
<td>T1</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>Attribute grading level</td>
</tr>
<tr>
<td>T1</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>Weight attached</td>
</tr>
<tr>
<td>T1</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

Table 7.33 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO ADVANCED SEARCH QUERIES

The findings of the present study with regard to advanced queries partly confirm the findings of Enríquez Raído (2011), who in her study of six participants observed that two of them used advanced queries, although most of these queries were attributed to one participant. They are, however, in contrast to Valli’s findings for concordance users whose search patterns resembled web searches and who “submit[ted] short queries and hardly use[d] advanced features” (Valli, 2012:7). However, the present study is not concerned so much with whether advanced searches are performed by professional translators or not. Rather, it seeks to find out whether advanced queries are used by a certain type of translator and are therefore a part of the research style topography. It is interesting to note that translators who used the most advanced searches were also observed to use planning as part of their research strategy (see section 7.1.2.5.). This would point to the fact that they are more methodical than those who do not use advanced search techniques. Interestingly, Enríquez Raído (2011) links the use of advanced queries with a more deep and sophisticated style of research, whereas in the present sample, three of the four most prolific users of advanced queries were also classified as having ‘strong presence’ of shallow searches. This is particularly true of T3, whose 66% of searches (26 research episodes) were classified as shallow as they were mostly very brief (two-three seconds) consultations in bilingual dictionaries.

7.1.2.9. Repetitive behaviour

Repetitive behaviour refers to repeating the search pattern consistently using the same set of resources in the majority of research episodes. It was observed that while many translators use a variety of resources to initiate and carry out their searches or choose different resources depending on the perceived need, some translators were observed to use the same resources more repetitively than others.

Due to the complexity of the data, the repetitive behaviour was only observed in relation to the initial search action, i.e. the first resource that was accessed in the research episode. This behaviour was considered repetitive if it occurred in over 50% of the instances, in this case, initial actions of research episodes. Figure 7.3 below shows that a total of nine translators were observed to start their searches with the same resource in over 50% of their research episodes. The most repetitive translator was T9, who started 96% of her searches with Google search engine, followed by T5, who initiated seven out of eight research episodes with Google Translate. Five translators were found to be moderately repetitive and two displayed weaker repetitive behaviour. Other participants who repetitively used Google were T4, T12 and
T16 and the remaining four translators (T3, T8, T13 & T14) initiated most of their searches with dictionaries.

<table>
<thead>
<tr>
<th>Translator</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of observed instances of repetitive use of a resource in initial search actions</td>
<td>n/a</td>
<td>n/a</td>
<td>28</td>
<td>6</td>
<td>7</td>
<td>n/a</td>
<td>35</td>
<td>22</td>
<td>n/a</td>
<td>4</td>
<td>16</td>
<td>14</td>
<td>n/a</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of all initial actions</td>
<td>0%</td>
<td>0%</td>
<td>69%</td>
<td>54%</td>
<td>87%</td>
<td>0%</td>
<td>41%</td>
<td>85%</td>
<td>95%</td>
<td>0%</td>
<td>0%</td>
<td>80%</td>
<td>80%</td>
<td>77%</td>
<td>0%</td>
<td>62%</td>
</tr>
<tr>
<td>Attribute grading label</td>
<td>W</td>
<td>W</td>
<td>MP</td>
<td>WP</td>
<td>SP</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>MP</td>
<td>SP</td>
<td>A</td>
<td>A</td>
<td>MP</td>
<td>MP</td>
<td>MP</td>
<td>A</td>
</tr>
<tr>
<td>Weight attached</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7.3 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO REPETITIVE BEHAVIOUR

Repetitive behaviour can be associated with an economising rather than exploratory type of behaviour; however, it has to be noted that two of the participants with moderately present repetitive behaviour translated from languages which may not have as wide a choice of resources to choose from, namely Hungarian and Indonesian.

7.1.2.10. Strategic behaviour

Strategic behaviour is characterised by additionally observed ‘tactics’ for finding the required information. These tactics relate particularly to a different way of dealing with word and world knowledge based research needs by employing a different chain of resources for each type of knowledge. Strategic behaviour differs from research strategies discussed in section 6.1.3. While research strategy only considers the initial action, i.e. starting from the known (a known resource) vs. starting from the unknown (search engine), strategic behaviour represents a more complex ‘chain reaction’ in response to a research need. A total of four translators were observed to use this kind of strategic behaviour, with T7 using these tactics most consistently and T1, T4 and T9 using clear tactics, but less consistently (see Table 7.34 below).

Table 7.34 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO STRATEGIC BEHAVIOUR

The four participants’ strategic behaviour will be briefly described below.

T1’s strategies related to employing different tactics to seek world and word knowledge. For the former, each research episode started with Google, followed by knowledge-based resources such as Wikipedia, and
for the latter, T1 used a wide range of resources, depending on what she was looking for. For example, in the audio commentary she said that for sentences she looks in Linguee and for words in other resources such as ProZ and WordReference.

50% of T4’s research episodes in which word knowledge was sought were initiated using Boolean operators for the exact phrase and the remaining five research episodes were all non-routine and started with a direct address search (URL), a web page, Linguee and a WordReference bilingual dictionary. In one of those five, a simple Google query instead of the exact phrase was used despite looking for the phrase ‘avoid the haircut’. In another instance, the exact phrase operator was overused for a single word ‘Bitcoin’. This would imply that T4 displays strategic behaviour, but those strategies are not always sensibly applied.

T7 was found to have clear routines and strategies throughout the task. Two types of tactics were observed - one associated with looking for word knowledge (using mostly ProZ Glossaries or Linguee as the initial port of call) and a three-step strategy for world knowledge. In the latter strategy, T7 first looked for a definition in TermWiki, then for an equivalent in other resources such as ProZ or Linguee and concluded by confirming the findings through Google search. This strategy was observed for key terms such as ‘fiat money’, ‘haircut’ and ‘crypto-currency’.

T9 consistently applied a hybrid search strategy using Google, based on formulating the query to include the source word, the language and the resource she had in mind (discussed in 6.1.3.). This query technique allowed her to get as much information as possible from the snippets to cover her research needs, in most cases related to word knowledge.

7.1.2.11. Drive for perfection

Drive for perfection in this study refers to a situation in which the process of arriving at a solution is subject to many revisions or where a significant amount of effort, whether in terms of time or resources, is expended to meet a research need. This can be best observed in T1, a Spanish native speaker researching the word ‘falter’, which is described in detail below.

1. Comes to the word ‘falter’ and puts a tentative solution down (‘desaparece’) saying “this is more or less what I think it is”.
2. Goes to ProZ Glossaries, searches for ‘falter’ but says “I don’t like either one of those, so I go to the next one”.
3. Goes to WordReference, finds a word ‘flaquear’, says “this is more or less what I am looking for”.
4. Uses ‘flaquear’ to search for synonyms in WordReference, but says “I don’t like any of those”.
5. Decides to go to Reverso, saying that “sometimes the others [i.e. other resources] have options that you don’t find in the first one”. Types in ‘falter’, but does not find anything suitable.
6. Looks through the resources in her ‘favourites’ tab, chooses Collins English-Spanish dictionary and searches for ‘falter’, finds the word ‘fallar’, saying “well, this one is the best one I have seen”.
7. Puts ‘fallar’ in the search box of Collins English Thesaurus. However, no search is performed.
8. Goes to the monolingual dictionary Real Academia Española and puts query ‘fallar’ in the search box. Finds the word ‘resquebrajar’ and says ‘it is a little bit what I’m looking for’.
10. Clicks on the first result which takes her to WordReference, where she looks at the synonym results, but does not comment on what she finds.
11. Goes back to the text, changes ‘desaperece’ to ‘empieza a desaparecer’.
12. Switches back to WordReference, puts in query ‘desaparecer’ in the search box. Looks at the results and says ‘I found it’.
13. Puts the word ‘desvanecer’ in the text.

This procedure related to seeking word knowledge took 11 steps, 8 specific resources and around 6 minutes. The research path looked as follows:

ProZ Glossaries → WordReference bilingual dictionary → WordReference synonyms → Reverso → Collins English-Spanish bilingual dictionary → Collins English thesaurus → monolingual dictionary Real Academia Española → Google → WordReference synonyms → WordReference synonyms → WordReference bilingual dictionary

This example not only shows the huge research volume, time, intensity and diversity of the translator’s research, but it points to T1’s unyielding determination to find the word that best fits with her vision. At the other end of the spectrum are translators who produce ‘good enough’, ‘fit for purpose’ translations and do not dwell on finding what they would consider a perfect or near perfect solution. An example is T14, who, answering the question ‘Do you feel that the quality of the translation was the same as if it was for a real client or do you think you would have spent a bit longer on the task in a real-life scenario?’, admitted that he proceeded as if it was a real assignment from a client, but also said that “I usually don’t spend too much time on translations”.

Table 7.35 below shows that four translators (T1, T6, T11 and T15) exhibited behaviour in which extensive efforts were made in the pursuit of a solution which was acceptable to them. Drive for perfection could be perceived as inefficiency in using resources to find an acceptable solution in as few steps as possible, but translators who displayed this feature appeared to know their resources very well. This behaviour can be associated with explorative research patterns and, strikingly, all these translators, apart from T15, also exhibited other explorative behaviours such as squirrelling or affective assessment.

Table 7.35 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO DRIVE FOR PERFECTION

---

222 Point 1 and 13 are not considered to be steps, therefore there are 13 points describing T1’s research, which took 11 steps.
223 It has to be reiterated at this point that this study is not concerned with the quality of the translations produced and the success rate of the queries. Therefore, it is not possible to assess whether or not the increased efforts have actually resulted in a solution of superior quality.
7.1.2.12. Affective assessment

Affective assessment refers to a behaviour whereby emotions are expressed during the research process and accompany the choice of a solution. An emotional assessment does not necessarily mean that choices are made irrationally. Rather, as Hansen (2005) points out, it is an additional element of the decision-making process. Certain expressions were used in the translators’ verbalisations to indicate either their like or dislike of a particular solution, for example, ‘wow, that’s cool’, ‘nice’, ‘I like that a lot’, ‘I like this – this is the word I was looking for’ and ‘I’m not happy with this’, it doesn’t sound good – it’s just a matter of preference’.

Three translators used emotional expressions during the task, and they happened to be the same translators who displayed squirrelling behaviour, i.e. T1, T6 and T11 (see Table 7.31). Therefore, emotional assessment can also be associated with explorative research behaviour patterns in the present sample. It has to be said, however, that this observation might have been an artefact of the methodological factors, i.e. the request to produce an audio commentary during the task.

<table>
<thead>
<tr>
<th>No. of observed instances of affective assessment</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of observed instances of affective assessment</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attribute grading label</td>
<td>SP</td>
<td>Δ</td>
<td>Δ</td>
<td>Δ</td>
<td>W</td>
<td>Δ</td>
<td>Δ</td>
<td>Δ</td>
<td>Δ</td>
<td>Δ</td>
<td>Δ</td>
<td>Δ</td>
<td>Δ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight attached</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range span</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.36 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO AFFECTIVE ASSESSMENT

7.1.2.13. High information retention

The level of information retention was self-declared by participants in the post-task questionnaire. Participants were asked to rate the level of their general retentive memory on a scale of 1-5, one being the lowest and 5 being the highest. They were asked whether they tended to retain information and knowledge with ease, e.g. facts, dates etc. Some data is missing or could not be established from the answers given (hence question marks in the Table 7.37); nevertheless, seven translators indicated a mid-range value (3) and seven rated the level of their retention as high (4-5). High retentive memory was mostly found to characterise translators whose volume of research was on the lower end of the scale. No conclusions could be made with regard to low retention as only one translator declared having a low retentive memory.

<table>
<thead>
<tr>
<th>Translators</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
<th>T16</th>
</tr>
</thead>
<tbody>
<tr>
<td>High information retention</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 7.37 CLASSIFICATION OF INDIVIDUAL TRANSLATORS ACCORDING TO HIGH INFORMATION RETENTION
7.1.2.14. Attitude towards technology

Attitude towards technology is an important factor in determining what technological solutions, including external resources, are employed by translators in their daily work. It has been pointed out in a previous publication that “the more readily people adopt new technologies and follow technological developments, the more tools they use and the greater the variety of software they tend to explore” (Gough, 2011:204). The same can be said about external resources, which form a part of translation technology. One of the key findings of Gough’s study of the adoption of web 2.0 technologies by translators is that “the underlying factor determining translators’ awareness, perception and use of tools and processes is their attitude towards adopting new technologies” (Gough, 2011:213). Massey and Ehrensberger-Dow also draw attention to the possible lack of awareness on the part of translators of the availability of online resources, coupled with the fact, as noted earlier, that some of them could not differentiate between online and electronic resources such as CD or DVD (Massey & Ehrensberger-Dow, 2011b:2003). In the present study, these findings are being taken further to see if attitude towards technology can have an impact on translators’ research style.

Translators taking part in the Main Study were asked in the profile questionnaire ‘How would you classify yourself with regard to adopting new technologies?’ with the following choice of answers:
- innovator and early adopter (looking for innovative solutions and picking up new technologies as soon as they emerge)
- fast follower (careful attitude but accepting change more quickly than the average)
- late majority (sceptical attitude and using new technologies when the majority are using them)
- traditionalist (only accepting innovative technologies when they have become commonplace or traditional)

As illustrated in Table 7.38 five translators classified themselves as innovators (T2, T3, T4, T9 & T10), eight as fast followers (T1, T6, T7, T11, T12, T14, T15, T16) and three as late majority (T5, T8 and T13). No-one identified themselves as a traditionalist, but interestingly, all three cases of late majority were associated with the understated research style, in which translators are on the lower end of the scale in terms of research volume, time, intensity and diversity. This would confirm earlier findings that attitude towards technology has a bearing on the use of technology and is a relevant factor to consider when discussing translation-oriented research styles (Gough, 2011).

<table>
<thead>
<tr>
<th>Attitude towards Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovator/early adopter</td>
</tr>
<tr>
<td>Innovator/early adopter</td>
</tr>
<tr>
<td>Innovator/early adopter</td>
</tr>
<tr>
<td>Late majority</td>
</tr>
<tr>
<td>Late majority</td>
</tr>
<tr>
<td>Innovator/early adopter</td>
</tr>
<tr>
<td>Innovator/early adopter</td>
</tr>
<tr>
<td>Fast follower</td>
</tr>
<tr>
<td>Fast follower</td>
</tr>
<tr>
<td>Late majority</td>
</tr>
<tr>
<td>Late majority</td>
</tr>
<tr>
<td>Fast follower</td>
</tr>
<tr>
<td>Fast follower</td>
</tr>
<tr>
<td>Late majority</td>
</tr>
<tr>
<td>Late majority</td>
</tr>
</tbody>
</table>

Table 7.38 INDIVIDUAL TRANSLATORS’ ATTITUDE TOWARDS TECHNOLOGY

7.1.2.15. Secondary categories summary

Based on the TTRS grid (See Figure 7.2) where the initial position of individual translators was established, the secondary categories were clustered based on the most commonly displayed secondary categories of the translators located in each of the five segments of the TTRS grid. For example, if most translators in the Economical segment of the TTRS grid displayed shallow searches, then it was assumed that shallow searches are one of the attributes of the Economical Style. The presence of secondary categories was assessed for each individual translator, resulting in either the strengthening of the existing position of a
translator who was already in a particular segment of the TTRS grid or in the shifting of the borderline translators to one of the neighbouring areas. Figure 7.4 below illustrates a ‘cloud’ of attributes which were associated with a particular segment of the TTRS grid based on the secondary categories identified in the unambiguous cases (i.e. excluding the borderline cases).

Figure 7.4 A CLOUD OF SECONDARY CATEGORIES ATTRIBUTABLE TO PARTICULAR SEGMENTS OF THE TTRS GRID

Based on the primary and the secondary categories, the final position of each individual translator was established. Figure 7.5 below illustrates how each of the participants’ position was either strengthened or altered according to the combination of both primary and secondary categories.

Figure 7.5 TTRS GRIDS SHOWING REPOSITIONING OF TRANSLATORS DUE TO THE INFLUENCE OF SECONDARY CATEGORIES

227
7.2. The Typology of Translation Research Styles (TTRS)

Based on the research behaviour of the sixteen translators observed in the Main Study of the present study, the following five research styles are proposed: the Prolific Translator, the Explorer, the Methodical Translator, the Economical Translator and the Understated Translator. Figure 7.6 below shows the final classification of the sixteen translators according to their research style.

Although taxonomies\textsuperscript{224} based on small samples can be seen as simplistic representations of complex phenomena (Bailey, 1994:15), Bailey convincingly argues that they can be very effective in bringing order out of chaos (Bailey, 1994:33); “they can”, Bailey continues, “transform the complexity of apparently eclectic congeries of diverse cases into well-ordered sets of a few rather homogeneous types, clearly situated in a property space of a few important dimensions” (ibid.). In the present study, this eclectic mix of professional translators was organised into groups on the basis of the similarity of the various aspects of their online research behaviour, but they by no means exhibit all behaviour assigned to the category they belong to. Below, each research style is described in terms of the shared aspects of the research behaviour of its members, followed by a description of the most prototypical members of the categories who are described as ‘cases’.

\textsuperscript{224} Bailey refers to taxonomies; however, it can be argued that this is also true for typologies.
7.2.1. The Explorer

Explorers enjoy researching and prioritise what they see as perfection over efficiency. For the Explorer, the research process is not just a means of satisfying their research needs but is a journey of discovery in itself. The main feature of this type of researcher is the considerable effort they exert in their research activities. Explorers use many resources and they research many words, terms and expressions, making sure that all words, even the lower risk ones (Pym, 2009:online), have been checked and confirmed. They use a broad variety of resources to look for ideas and inspirations and are keen to understand the nuances rather than settle for shallow comprehension. They like to keep abreast of developments, although they are not necessarily at the forefront of technological developments and are normally fast followers. Their research is characterised by deep searches and a relaxed pace. Many of their research episodes are short as Explorers like to check and double check their solutions a lot, as well as refine their choices. However, due to their inquisitive nature, Explorers can also have some very long research episodes, both in terms of the number of steps and the time taken to find a solution. Therefore, their research intensity overall is medium. Their search paths are often meandering and branch out, and sometimes they like to continue with their research for longer just to satisfy their curiosity or deepen their subject knowledge. Because they spend so much time looking for the perfect or near-perfect solution, they like to record these findings, whether in their termbases or glossaries. They will also be very likely to engage in communities and cast a vote for a solution in community-based resources. They can be very organised in their approach to research, but their research behaviour can appear rather chaotic due to the lack of routines. This could be due to their more creative style which favours an individual approach to each translation problem, rather than following pre-established research paths.

The prototypical Explorer in this sample, although not an ‘ideal’ or ‘pure’ type in the sense proposed by Weber (1949), is T1. T1 is a sworn-certified English to Spanish translator and interpreter aged 50-59, with over thirty one years of experience in many fields, including healthcare/medical, security/military, legal, finance/business and literary. In the post-task questionnaire, T1 confirmed that although she had no previous experience with digital currency, she was “highly experienced in financial translation” and was comfortable with the text. She reported that she approached the task as if it were a real assignment. However, with regard to the quality, T1 said she “would have spent a lot more time on editing” and she “did not complete many of the final quality assurance steps” she normally does.

Gox’ and ‘purport’. To research all these units she used the following nineteen resources: a monolingual dictionary (Real Academia Española), three bilingual dictionaries (Collins, WordReference, Reverso and Word Magic), three dictionaries of synonyms (Sinonymo.es, Synonyms in Word and WordReference synonyms), two glossaries (ProZ and own glossary), one thesaurus (Collins), one discussion forum (ProZ), one search engine (Google), one knowledge-based resource (Wikipedia), one concordancer (Linguee) and four web pages. Due to the high number of RUs and high resource volume, her overall research volume research was Generous. As can be seen, T1 not only used many specific resources, but also a wide variety of types and used them frequently (100 instances of resource use in the task), which makes her a very diverse translator. Her total research time was 1h 0’ 58’ which was by far the longest recorded time for the task. However, her Internet connection was slow, which had an impact on the length of the process, but given that in the post-task questionnaire she said she only got to 70% with the quality assurance, it could safely be said that T1’s research time was Generous.

Due to the extensive use of resources, T1’s research patterns were characterised by many breaks in the flow of the translation, with frequent ‘flicking’ between the texts and resources, although the actual duration of the research episodes was relatively long. The intensity of T1 appears to be moderate due to the balance between the number of long, medium and short, one-step consultations. Taking into account all this research activity, T1 was very organised, with all resources neatly arranged in a resource tab. She also used a colour-coding system to help her remember which parts of the translated text needed to be revisited. She kept a record of the solutions she found and often saved them in her personal glossary. T1 also contributed to community-based resources such as ProZ Glossaries and was observed to cast a vote on community-based sites whenever she liked a solution she had found. T1 followed mostly straight paths in her research and moved in a sequential manner, but no repetitive behaviour was observed within the sequences. Instead, T1 tended to rely on strategies which she has developed to find the type of knowledge she requires. Snippet viewing is another technique used by T1 in order to maximise her research experience. T1 knows her resources well and what to expect from them and uses them to “find a better or a more specific term” in her often lengthy pursuit of the perfect solution (see example in 7.1.2.11.), which is often accompanied by an emotional assessment of the solution.

7.2.2. The Prolific Translator

The Prolific translator is in many ways similar to the Explorer. With a generous volume of research, the prolific translator enjoys the discovery process and the creative rather than a routinised way of finding information. The wide range of specific resources and types of resources used by Prolific Translators, as well as the high frequency of resource use are also evidence of their high research volume. However, Prolific Translators process their high RU volume much faster than Explorers, which is expressed in their high research intensity. Their research is characterised by research episodes that are moderate in length and comprise a moderate number of research steps, but these episodes are
shorter than the Explorer’s thus resulting in moderate to rapid research pace. A prototypical Prolific Translator would, therefore, process a great number of RUs using a high number of varied resources in a relatively short time. Similarly to the Explorer, they would have a tendency to go deeper in their search, but the depth would not be related to the time expended but to the volume of material consulted. Prolific Translators might exhibit some meandering research paths and squirrelling behaviour, but in order to consult vast amounts of resources in a short period of time, they will mostly use techniques such as parallel search or snippet viewing to maximise the recall and consequently, exposure to potential solutions. They might also use machine translation as a fast way of accessing potential solutions to their translation problems.

The dataset did not reveal a clear Prolific Translator. The closest to this style was T11, who is an English into Polish translator aged 30-39, with five years of experience. T11 specialises in law, business and administration and has a BA in Applied Translation and an MA in Translation. She had no previous experience with digital currency, but had translated business texts and did not feel completely out of her comfort zone. The text was “approached very seriously, as if it were a text to be published”, but she would have spent a little longer in a real life scenario – “at least one review more”. T11 reports that having to comment slowed her down considerably, but her interaction with online resources resembled her natural way of researching.

T11 researched the following thirty one research units: ‘digital currency’, ‘rise’, ‘Bitcoin’, ‘WIR bank in Switzerland’, ‘WIR bank currency’, ‘Ithaca HOURS system’, ‘commodities based currencies’, ‘gold standard’, ‘portfolio’, ‘longevity’, ‘disruptive’, ‘mainstream’, ‘fiat money’, ‘money transfer systems’, ‘banking systems’, Liberty Reserve’s LR currency’, ‘hard currency’, ‘according to’, ‘smartphone’, ‘haircut’, ‘Cypriot bank’, ‘open source’, ‘crypto-currency’, ‘peer-to-peer’, ‘traded online’, ‘subdivided down to eight decimal places’, ‘satoshi’, ‘mining’, ‘currency exchange’, ‘Mt. Gox’, ‘to manage bitcoin traffic’. To research these units, T11 used nine types of resources and twenty eight specific resources which were accessed on 138 occasions. The following resources were used: a monolingual dictionary (businessdictionary.com), a bilingual dictionary (megasłownik.pl), a term bank (IATE), a search engine (Google), two knowledge-based resources (Wikipedia and a Wiki-type resource), eighteen web pages, machine translation (Google Translate), two digital documents (a pdf and a ppt) and call an expert, which was initiated, but not carried through. Her total recording time was second longest (1h 58’23’’), of which translation took 1h 4’57’’ and research 30’39’’.

T11’s research was characterised by a prolific use of parallel texts and frequent consultations in Google Translate. T11’s use of dictionaries and other termino-lexicographic resources was limited; instead, she gathered information from the many parallel texts she had read in advance and during the translation task. Another frequently used sequence was to consult Google translate, often out of curiosity or to find collocations, by putting words, phrases and even sentences to see “how it [GT] would translate it”, which would then normally be followed by a search in Google and Wikipedia or in a parallel text. T11 displays characteristics of an Explorer, such as meandering, squirrelling (by adding words to her Word dictionary during the spellchecker) and affective assessment (especially with regard to GT output). However, she also had a strong presence of parallel searching, with the highest score of twenty one instances. T11’s modus operandi when searching through Google results was to click on the links of possible interest and open them in separate windows, sometimes three or four from a single page of Google results. Some of these opened tabs were never consulted; however, most of them were. T11’s browser was always full of opened tabs (on one occasion, fourteen open tabs were counted), and she had to perform separate ‘tab closing’
actions during the task. Although T11 was not initially classified as a Prolific Translator as the primary categories grid positioned her right on the border of the Explorer and the Prolific, the additional strong presence of behaviours indicating a fast processing of the high research volume justifies the classification of her as a Prolific Translator.

7.2.3. The Methodical Translator

The Methodical Translator carries out research in a systematic rather than a discovery mode and the main characteristic of this type is moderation. This moderation is often self-imposed by means of self-organisation and planning in the initial phase or by having clear strategies for dealing with various types of translation problems. Advanced search is also often used as part of the planned strategies. The research behaviour of the Methodical type is characterised by moderate research volume and research time although sometimes this may fall towards the lower end of the scale. The diversity of the resources will also be mid-range and the searches moderate in depth. Overall, the Methodical Translator tends to display research behaviour that centres on the middle ranges. Search paths are mostly straight and this type looks up resources in a sequential manner, avoiding excessive numbers of open tabs in their browser. Rather interestingly, all four Methodical Translators self-declared that they were innovators/early adopters.

The most ‘prototypical’ translator in this category is T4, who is an English to Spanish translator aged below 29, specialising in medical and audiovisual translation. She holds a BA in Translation, an MA in Audiovisual Translation and an MA in Medical Translation. T4’s research volume was at the lower end of the spectrum, and she researched the following eleven units: ‘WIR bank in Switzerland’, ‘Ithaca HOURS System’, ‘commodities’, ‘gold standard’, ‘Bitcoin’, ‘fiat money’, ‘Liberty Reserve’s LR currency’, ‘avoid the haircut’, ‘open source crypto-currency’, ‘satoshi’ and ‘Mt. Gox’. The specific resources she used in the task included: a bilingual dictionary (WordReference), a search engine (Google), a knowledge-based resource (Wikipedia), a concordancer (Linguee), five web pages and two online pdf documents. T4 had no experience with the subject of digital currency, however, she “was still comfortable with the text” as it “did not present too many problems”. She confirmed that she proceeded with the task as if it were a real assignment, but admitted that in a real life scenario she would have reviewed it again before delivering. She also declared that her interaction with online resources closely resembled the way she normally interacts with them. T4 is a typical planner. She read the text systematically and highlighted thirteen units saying these were the terms she would be researching as she goes. She then methodically proceeded with her translation and did not research anything else but those highlighted research units. In fact, she admitted in the audiocommentary that she did not need to research two of the highlighted terms. While translating, she displayed strategic behaviour in the way she repetitively formulated her queries in the same way and used advanced Google searches in the form of exact phrases as initial search actions. T4 proceeded in a straight
path and a sequential manner and did not diverge from anything that was not planned at the beginning of
the translation process.

7.2.4. The Economical Translator

Economical Translators are characterised by a pragmatic attitude towards research. The key
factors driving their research are time and efficiency and their research effort is inversely
proportional to that of the Explorer. They are fast-paced and only research essential items in the
source text; therefore there is little checking or confirming behaviour in their searches, which are
mostly shallow, composed of few steps and follow straight paths. This style of research is
characterised by routine behaviour in which searches are frequently initiated using the same
resource or the same pattern of resource use is repeated in many of the searches, sometimes
independently of the type of knowledge sought.

The Economical Translators have high retentive memory, which might contribute to their low research
needs. They might use machine translation to aid their research process or speed up the translation
process. They are often end-revisers and five out of eight identified Economical Translators performed
rapid translation, with long end revision. To summarise, the Economical Translator is a fast researcher for
whom research is not an activity they will delve into deeply.

The prototypical Economical Translator in this category is T16, who is an English to Spanish translator, aged
30-39, with six to ten years of experience, specialising in business and marketing, IT, healthcare and general
translations. She has a BA in Translation and Foreign Languages and an MA in Translation and Interpreting.
In the post-task questionnaire, T16 confirmed that she was familiar with the subject matter from articles
she had read and was comfortable with the text. She also confirmed that she approached the task as if it
were for a real client and said that the “time spent was pretty close to what I would have spent in a real life
situation”, although she admitted that she might have spent “a bit longer in a real life scenario”. T16
confirmed that the experience of the task was “pretty close to a real experience” and that she tried “to go
about it as I would normally with any given project”.

In the translation task, T16 was observed to investigate the following eight research units: ‘Ithaca Hours’,
eleven specific resources: a bilingual dictionary (Real Academia Española), a monolingual dictionary (OED),
a search engine (Google), a concordancer (Linguee), a knowledge-based resource (Wikipedia), five web
pages and an online pdf document and accessed them on twenty four occasions. Her timings were the
shortest of all participating translators. Her total recording time was 28’33”, translation time 16’44” and

225 Four out of eight identified Economical translators said in the post-task questionnaire that their retentive memory
is high (5 on a 1-5 scale, 5 being the highest).
her research time was the shortest of all participants\textsuperscript{226} at 6’12”. She performed her searches fast, getting only the information she required and her searches were shallow and often cursory.

T16 admitted in her post-task questionnaire that she is neither a research-intense nor a methodical type and that although she generally tries to acquire as much information as possible beforehand, “ultimately, there will be problems during the translation process that [she] had not foreseen […], so there is no point spending a lot of time trying to foresee those”. Lastly, in the audio-commentary, T16 talked about efficiency and how, in view of the fast way the industry works, she cannot afford to do what she was taught at school (for example read the text very well before translating) and so she feels the need to adapt to a different way of dealing with text, i.e. faster and more efficient. T16 was initially positioned on the border with the Methodical style, but due to the strong influence of secondary categories, her position was adjusted to fit the Economical style.

7.2.5. The Understated Translator

The main characteristic of Understated Translators could be summed up by the ‘less is more’ maxim. Frugality is the main feature of this type of translator; they will only research items that are absolutely necessary to achieving the desired quality of their translation, and tend to resort to the use of internal resources to solve lower risk problems (Pym, 2010:online). Therefore, they research few words or terms and have at their disposal only a handful of resources which they access routinely to satisfy their research needs. Although they use few resources, the ones they do use are of different types, providing a variety to choose from, depending on the research need. Their searches might be sporadic, but Understated Translators do take their time to research them well, their research being marked by a high proportion of deep searches and indirect, serendipitous research. Translators representing this research style will be driven by perfection, but rather than progressing in an exploratory manner, they will use advanced search queries or certain established strategies to make sure their research needs are fully satisfied. To sum up, Understated Translators research little using a handful of reliable resources, but their research is thorough and exhaustive.

A good example of an Understated Translator is T7, who is an English to Brazilian Portuguese translator in his fifties with eleven to twenty years of experience. He specialises in IT, ERP software and life sciences. He has a degree in engineering but attended translation courses. T7 was comfortable with the text and confirmed that he proceeded as if it were a real assignment, apart from not running his AQ tool. He also reported that the audio commentary did not have much impact on his performance, other than slowing him “a tiny bit” and that his interaction with online resources resembled his normal research process.

\textsuperscript{226} Apart from T2, who did not carry out any research.

234
T7 researched the following eleven research units: ‘commodities’, ‘savings’, ‘portfolio’, ‘disruptive’, ‘mainstream’, ‘fiat money’, ‘wild swings’, ‘haircut’, ‘crypto-currency’, ‘peer-to-peer’ and ‘purport’, using five specific resources, each of a different type, accessing them on thirty one occasions during the task. The resources used were as follows: an orthographic dictionary (Orthographic Vocabulary of the Portuguese Language), a Glossary (ProZ), a term bank (TermWiki), a search engine (Google) and a concordancer (Linguee). His total recording time was 46’02’’, translation time 24’38’’ and his research time 16’51’’. His research was characterised by measured to relaxed pace and clear-cut strategies for dealing with word and world knowledge. When he encountered a problem related to word knowledge, he used ProZ Glossaries or Linguee. However, when he came across a world knowledge related problem, he had a three-step strategy (see section 7.1.2.10.), which he used consistently.

7.2.6. TTRS Summary

As illustrated in the previous sections, half of the sample clusters around the Economical style. The second most populated category is the Methodical research style, with 25% of the sample displaying characteristics of this style. The other styles have fewer representatives and it important to remember that the Prolific style did not have a clear member after the initial positioning. Two translators (T3 and T6) were problematic with regard to the final positioning. T3 displayed characteristics of the Economical as well as the Methodical style but scored slightly higher on the Economical one and was therefore moved accordingly. T6, although initially positioned as the Understated style, displayed strong characteristics of the Explorer. She was therefore positioned very close to the Explorer border to reflect this.

Below is a summary of the most prominent characteristics of each research style which includes the primary characteristics, secondary attributes and additional descriptors for each style.

![Figure 7.7 SUMMARY OF CHARACTERISTICS OF RESEARCH STYLES](image-url)

At this point, it is important to explore some tentative connections with the studies of Carl et al. (2011) and Dragsted & Carl (2013) who proposed a taxonomy of human translation styles. In the present study, both Understated Translators and six out of eight Economical Translators were either head starters or quick
planners and five were end revisers. It could therefore be concluded that the translators on the lower end of the volume axis (researching fewer terms, using fewer resources of lesser diversity etc.) prefer to start their translation without much planning and then revise heavily at the end. The Methodical Translators, on the other hand, displayed a mix of skimming and systematic orientation and mostly constant or online revision, which is also the case for the Explorative and the Prolific types at the revision phase. This would suggest that the more research carried out in terms of the volume of RUs and the resources, the more planning or explorative tendencies displayed and the more evenly spread the revision throughout the task. However, when comparing the local vs. global strategies proposed by Dragsted & Carl (2013:149) with these results, there seems to be an incompatibility. In Dragsted & Carl’s study, the local strategy is characterised by head start and online revision while in the global planning strategy translators are seen to have a more systematic orientation and to revise at the end. It would appear that in the present study, those translators who had little initial orientation, i.e. the head starters, were also end revisers. Conversely, the planners were mostly observed to revise as they progressed through the task. Although these findings have to be treated with caution as they were not the primary goal of the present study and are based on incomplete data, they do present a stimulus for further investigation of the relationships between translation styles and research styles.

This overview of the five research types, although comprehensive, by no means constitutes an exhaustive or complete classification. Though many aspects of the styles are clear-cut, some still require refining and sounder methodological underpinning. Therefore, it is suggested that this typology should be regarded as an initial conceptual construct rather than a definitive solution, encouraging further research in this area.

7.3. The influence of the demographic and professional status variables on RTUT and TTRS

The main tenet of the present work is that the professional translators who participated in the study seem to display various research styles; preliminary indications are that age, gender, experience or knowledge domain do not seem to play a clearly identifiable role in influencing these ‘styles’. A shallow analysis of the influence of the above mentioned variables on the components of TTRS and RTUT as well as on the final classifications of translators according to their resource and research styles is provided, suggesting that although some tentative associations between these variables and the typologies could be identified, given the sample size, they do not provide clear enough evidence to suggest that the resource or research styles depend on these variables. Nevertheless, this analysis will be considered here in order to establish groundwork for any future studies wishing to investigate these relationships further.

---

227 The information about the translation phases was gathered in the profile questionnaires and then observed in the screen recordings, i.e. the participants were explicitly asked to classify themselves in terms of all three phases based on Carl et al.’s taxonomy and then the actual behaviour in terms of these phases were noted from the screen recordings. In many cases the self-declared data matched the observed data. However, in some cases it either did not or one of the phases was missing in the recording. Therefore, when observational data was missing, only the self-declared data could be accounted for.
7.3.1. Length of experience

Experience has always been one of the main variables in determining how experts differ from novice translators. Much translation process research has been carried out comparing experts and novice translators, focusing on the differences between them in order to help develop translator competence. Whatever differences were found were mostly attributed to the level of expertise or to the length of experience. But the findings were often contradictory. Whilst some researchers presented evidence to suggest that more experienced translators use reference material less frequently than novice translators or that students rely more on bilingual than monolingual dictionaries (Jääskeläinen, 1996; Jensen, 1999), Livbjerg & Mees, 2003) others found the opposite to be true (Zheng, 2014). With regard to research activities it has been suggested that more experienced translators make fewer consultations than their less experienced colleagues (Zheng, 2014:19) and that they carry out deeper, more targeted and engaged research (Enríquez Raído, 2011) whilst less experienced ones struggle even with the use of the Internet (Dam-Jensen, 2012). However, opposite views can also be found in the recent literature (Massey & Ehrensberger-Dow, 2011b; Sales & Pinto, 2011), thus raising the question as to whether there are other factors at play that could account for these differences. Zheng (2014:12) rightly suggests that “consultation methods are related to multiple factors such as text style, time pressure and personal preferences”.

Therefore, whilst it is important to acknowledge that experience is a key differentiator between novice and experienced translators (however those differences are described), it has to be borne in mind that the differences can also be attributable to other factors. Table 7.39 below shows how RTUT and TTRS components relate to translators’ length of experience.

Table 7.39 THE INFLUENCE OF EXPERIENCE ON THE RTUT and TTRS COMPONENTS

<table>
<thead>
<tr>
<th>Translator</th>
<th>T11</th>
<th>T3</th>
<th>T4</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T2</th>
<th>T13</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of experience</td>
<td>3-5</td>
<td>6-10</td>
<td>6-10</td>
<td>6-10</td>
<td>6-10</td>
<td>6-10</td>
<td>11-20</td>
<td>11-20</td>
<td>11-20</td>
<td>11-20</td>
<td>21-30</td>
<td>21-30</td>
</tr>
<tr>
<td>Nature of resources</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>TL</td>
</tr>
<tr>
<td>Nature of resources frequency</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>M</td>
<td>TL</td>
<td></td>
</tr>
<tr>
<td>Research direction</td>
<td>M</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>TT</td>
<td>M</td>
<td>ST</td>
<td>M</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
</tr>
<tr>
<td>Research strategy</td>
<td>M</td>
<td>RI</td>
<td>TO</td>
<td>RI</td>
<td>M</td>
<td>RI</td>
<td>TO</td>
<td>RI</td>
<td>RI</td>
<td>RI</td>
<td>RI</td>
<td>RI</td>
</tr>
<tr>
<td>RTU volume</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Resource volume</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Research intensity</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Research diversity</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>Research time</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Research pace</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>M</td>
</tr>
</tbody>
</table>

TL - terminological resources user
CB - corpus-based resources user
M - mixed user
ST - source-text oriented user
TT - target-text oriented user
RI - research interests
RIU - research intensity
TUT - top-down user
BU - bottom-up user
L - low
M - medium
H - high

The more experience the:
- lower the RIU volume
- less resources used per RIU with greater experience
- lower diversity
- lower research time

Table 7.39 THE INFLUENCE OF EXPERIENCE ON THE RTUT and TTRS COMPONENTS
Table 7.40 below shows how RTUT and TTRS relate to translators’ length of experience.

Table 7.40 THE INFLUENCE OF EXPERIENCE ON RTUT AND TTRS

These tables show that there is no demonstrable evidence showing that the length of experience has a clear impact on the types of resources that are used and how they are accessed (RTUT), although there is some evidence to suggest that the length of experience was found to have some impact on the research style. For example, as illustrated in Table 7.39, on the whole, a higher proportion of the translators with the longest experience seem to have researched fewer items, used slightly fewer resources of lesser diversity and used them less frequently. With regard to time-related features, the more experienced translators spent slightly less time on research activities. However, looking at the breakdown into the real research time and the time spent on research in proportion to translation, it transpired that the times were not lower when measured in relative terms which would imply that translators spent less time on the task as a whole. Slightly longer research steps and episodes were also observed for the most experienced translators which resulted in slightly higher research pace. All these factors resulted in the clustering of the more experienced translators around the Economical style. Interestingly, all Methodical translators seemed to have less experience, between six to ten years (see Table 7.40). However, despite the fact that some links could be seen, they were rather weak; for example T1’s behaviour contradicted all of the points made above and T8 and T16, although much less experienced, displayed most of the characteristics of the more experienced translators.

7.3.2. Age

The second variable to be considered is age. The components of the RTUT and TTRS were cross-tabulated with the ages of the translators participating in the Main Study.
Table 7.41 below shows how the individual components of TTRS and RTUT relate to translators’ age.

<table>
<thead>
<tr>
<th>Translator</th>
<th>T4</th>
<th>T9</th>
<th>T11</th>
<th>T16</th>
<th>T15</th>
<th>T10</th>
<th>T12</th>
<th>T14</th>
<th>T13</th>
<th>T17</th>
<th>T18</th>
<th>T2</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of resources</td>
<td>CB</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>CB</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>CB</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>n/a</td>
</tr>
<tr>
<td>Nature of resources frequency</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
</tr>
<tr>
<td>Research direction</td>
<td>M</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>M</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>M</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
</tr>
<tr>
<td>Research strategy</td>
<td>TD</td>
<td>M</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
</tr>
<tr>
<td>RU volume</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>H</td>
</tr>
<tr>
<td>Resource volume</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Research intensity</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>Research diversity</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>Research time</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Research pace</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>H</td>
</tr>
</tbody>
</table>

Table 7.41 THE INFLUENCE OF AGE ON THE RTUT AND TTRS COMPONENTS

Table 7.42 below illustrates how RTUT and TTRS relate to translators’ age.

<table>
<thead>
<tr>
<th>Translator</th>
<th>T9</th>
<th>T4</th>
<th>T11</th>
<th>T16</th>
<th>T15</th>
<th>T10</th>
<th>T12</th>
<th>T14</th>
<th>T13</th>
<th>T17</th>
<th>T18</th>
<th>T2</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTUT</td>
<td>Mixed</td>
<td>PTF</td>
<td>PTF</td>
<td>Mixed</td>
<td>Mixed</td>
<td>DE</td>
<td>DE</td>
<td>DE</td>
<td>PTF</td>
<td>Mixed</td>
<td>DE</td>
<td>Mixed</td>
<td>n/a</td>
</tr>
<tr>
<td>TTRS</td>
<td>M</td>
<td>M</td>
<td>P</td>
<td>Ec</td>
<td>M</td>
<td>M</td>
<td>Ec</td>
<td>Ec</td>
<td>Ec</td>
<td>Ec</td>
<td>U</td>
<td>U</td>
<td>Ex</td>
</tr>
</tbody>
</table>

Table 7.42 THE INFLUENCE OF AGE ON RTUT AND TTRS

As before, it is only possible to make tentative associations. As with the length of experience, older translators researched slightly fewer RUs. Also, fewer resources in terms of type and diversity were found.
in the older translators’ toolboxes and they were used less frequently. The oldest translators had marginally longer research steps and research episodes resulting in higher research pace, which could indicate slightly more engaged and deeper research characteristics (see Table 7.41). Again, as in the case of experience, lower age seemed to characterise Methodical translators as they were all found amongst the younger cohort (see Table 7.42). As previously, these links should be treated with caution as many of these patterns are not consistent across all the participants and only the most general trends are highlighted.

7.3.3. Gender

The third variable against which the RTUT and TTRS components were examined is gender. Table 7.43 and Table 7.44 follow the pattern of the previous tables, showing first the influence of gender on the individual RTUT and TTRS components and then on the final resource and research styles.

Table 7.43 THE INFLUENCE OF GENDER ON THE RTUT AND TTRS COMPONENTS

<table>
<thead>
<tr>
<th>Translator</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>RTUT</td>
<td>M</td>
<td>M</td>
<td>CB</td>
<td>TL</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
</tr>
<tr>
<td>T1</td>
<td>n/a</td>
<td>M</td>
<td>M</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
</tr>
<tr>
<td>Research direction</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>M</td>
<td>ST</td>
<td>M</td>
<td>TT</td>
<td>n/a</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
</tr>
<tr>
<td>Research strategy</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
</tr>
<tr>
<td>T2</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Resource volume</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Research intensity</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Research diversity</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Research time</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Research pace</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
</tbody>
</table>

Table 7.44 THE INFLUENCE OF GENDER ON RTUT AND TTRS

Surprisingly, gender seems to suggest clearer if still tentative - given the sample size - association with resource and research styles. Male participants generally appeared to be more termino-lexicographically oriented, displayed ST-oriented dominance and used mostly top-down techniques. This resulted in three out of the five men taking part in the study being classified as Dictionary Enthusiasts. With regard to their research style, as shown in Table 7.43, men on average had slightly lower volumes of RUs and used fewer resources
per task and per RU. Their research time was also lower than that of most of the women taking part in the study. This resulted in four out of five men being classified as having an Economical style (see Table 7.44). Although these associations appear to be stronger than those related to experience and age, it should be noted that the sample was not equally composed of male and female participants and any apparent patterns therefore have to be treated with caution.

7.3.4. Familiarity with the domain

Familiarity with the business/economic/financial domains to which the assigned translation text belonged seems to have less influence on the RTUT and TTRS than intuition might indicate. Table 7.45 shows how familiarity with the domain relates to various components of resource and research styles, whilst Table 7.46 shows the overall influence of the familiarity with the domain on the resource and research styles.

<table>
<thead>
<tr>
<th>Translators</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of resources</td>
<td>M</td>
<td>n/a</td>
<td>T/L</td>
<td>M</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Nature of resources frequency</td>
<td>M</td>
<td>n/a</td>
<td>T/L</td>
<td>M</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>M</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>CB</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Research direction</td>
<td>M</td>
<td>n/a</td>
<td>ST</td>
<td>ST</td>
<td>M</td>
<td>T/L</td>
<td>M</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>ST</td>
<td>T/L</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Research strategy</td>
<td>BU</td>
<td>n/a</td>
<td>BU</td>
<td>M</td>
<td>TD</td>
<td>M</td>
<td>BU</td>
<td>TD</td>
<td>TD</td>
<td>BU</td>
<td>BU</td>
<td>BU</td>
<td>M</td>
<td>B</td>
<td>M</td>
</tr>
<tr>
<td>RU volume</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>Resource volume</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>M</td>
</tr>
<tr>
<td>Research intensity</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Research diversity</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Research time</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>Research pace</td>
<td>H</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.45 THE INFLUENCE OF DOMAIN FAMILIARITY ON THE RTUT AND TTRS COMPONENTS

Table 7.46 THE INFLUENCE OF DOMAIN FAMILIARITY ON RTUT AND TTRS

With regard to resource style it could be said that, with the exception of T5, as shown in Table 7.45, more of the translators with domain familiarity used a higher proportion of text-based resources, hence the higher proportion of Parallel Text Fans within this group. Some tentative links between the volume-related features of the TTRS were also noted. For example, research intensity appeared to be higher within the group familiar with the domain, which is seemingly counter-intuitive because it suggests that this group
needed to consult more resources for each translation problem than those who were not familiar with the domain. The same argument could be applied to the higher resource diversity observed within the group more familiar with the domain; however, this could also be linked to the greater use of text-based resources, and hence perhaps a higher number of websites being consulted. A higher proportion of the participants more familiar with the domain also had lower research times, but, as mentioned earlier, the overall time spent on the task was also lower and in terms of the proportions, there was no perceptible difference between the two groups. Although stronger associations between domain expertise and styles might have been expected, their weak nature could be linked to the fact that the text spanned several domains and related to the relatively new phenomenon of an alternative currency. Some translators, although experienced in the relevant domains, could have been unfamiliar with this concept. This was the case with T1, who, although very experienced, mentioned in the audio commentary that she had never heard of such a concept. T12, on the other hand, commented that he was very familiar with the subject and that some of his friends were actually farming (mining) Bitcoins.

As mentioned above, any relationships noted above need to be treated very cautiously; however they cannot be ignored. A deeper analysis is required in order to better understand these relationships, especially in view of the fact that some of the typology components are compound (see section 7.1.1.) and differences could be noted on the sub-category level. Nevertheless, it seems that there is still more variety within the studied group of professional translators than might initially have been assumed. Therefore, when looking at the patterns of online information behaviour of professional translators, any variables such as experience, age, gender or familiarity with the domain should be considered together with other factors influencing this behaviour such as resource style (RTUT) and research style (TTRS).
8. Summary of findings and discussion

The three research questions posed by this study led the investigation into a number of areas, focusing on the time spent on research activities, the types of resources used by professional translators and the patterns of the interactions with online resources during their research activities. In the process of the analysis, it turned out that some of these questions were easier to answer than others. Whilst it could be said that RQ1 and RQ2 were closed questions to which relatively straightforward answers were obtained, RQ3 was an open question and therefore invited a more comprehensive, multidimensional investigation. Because of this open nature of RQ3 and the broad conceptualisation of the possible outcomes as ‘patterns of interactions’, RQ3 yielded more findings and of a more varied nature. The summary and the discussion of all the findings will be presented below.

8.1. Time spent on research activities

As mentioned in section 5.2.1., despite various estimations of how much time translators spend on looking up information related to their translation work, research has failed to provide solid empirical evidence to support these claims. One of the main contributions of the present study to translation process knowledge is an attempt to empirically measure the time spent on research activities in a sample of sixteen professional translators, as set out in RQ1.

The findings of the present study suggest that on average translators spend a third of their time on research activities. The figures from both study components (including self-declared and observed data) show that in familiar domains, 21-30% of translation time is spent on research activities, whilst in an unfamiliar domain this time increases to 36-50%. If self-declared data is excluded, the figures indicate that 30% and 36% of the time is spent on research for familiar and unfamiliar domains respectively. Conflating the self-declared and the observed data from both study samples into one average revealed that translators spend 34% of their time on research activities during their translation assignments.

One of the findings from the Global Survey is that research activities reportedly take the same amount of time regardless of whether translators use technologies such as translation memory (TM) systems or terminology management systems (TMS). In the Main Study this was confirmed to some degree, although the use of TM and TMS was not extensive. Five out of sixteen translators used TM (2x Déjà Vu, 2x Wordfast Anywhere and SDL Trados), but only one participant explicitly said that she does not use TM in her translation work. Two translators used the MT option from within their CAT Tool and one participant used a pre-translate option despite claiming to have no ‘reliable’ TM to use. One translator was observed to actively use a termbase within Déjà Vu. Four out of five of the translators using TM had a low research time; however, it is not clear whether this was due to the use of TM as some of them said they did not have an appropriate TM database for this project and one translator started with an empty TM. Therefore, it is difficult to assess whether the use of TM had any impact on the research time in the present study as the text was short and non-repetitive. There is no doubt that a well-populated TM would have contributed to a reduction in research time; however, it was not possible to verify this in the present study.

228 With regard to the use of MT, the present research shows that research time does not seem to diminish greatly with the use of MT solutions although Daems et al.’s study shows otherwise (Daems et al., 2015). The two translators who worked in a post-editing mode had low
research times, both in real time and in proportion to translation activities. However, their times were not
the lowest within the sample as a few other translators had lower research times without using MT.

These findings have implications for technology development. Given that research activities take up such a
large proportion of translators’ time, it is surprising that most developments have been focused on
reducing the time associated with translation activities, whereas not much effort has been dedicated to
streamlining research activities and thus reducing the research time. One explanation could be that
developers assume that the leveraging of more content through TM and/or MT would lead to less need for
research. However, this assumption seems counter-intuitive. Admittedly, more words can be translated
with TM and MT, but at the same time proportionally more potential problems requiring research might
emerge precisely because of the increased volume of translated material.

A more plausible explanation for the lack of efforts invested in streamlining research activities could be
because these activities are not subject to remuneration in the per-word pay model currently most
dominant and therefore no gains from investment in this part of the process can be seen for those at the
end of the production chain. From the translator’s point of view, research activities could be linked to the
notion of ‘shadow work’, a concept introduced by Illich in 1981 to refer to the “form of unpaid work which
an industrial society demands as necessary complement to the production of goods and services” (Illich,
1981:100). Marcu (2015:online) argues that emerging technologies which do not yet provide consistent
quality, such as speech recognition and machine translation, force us to do ‘shadow work’, for example by
having to go through lengthy automated messages or by post-editing poor quality MT.

Carrying out translation-oriented research could also be classified as ‘shadow work’ because translators are
not paid for the time they devote to research. This was actually mentioned by one of the participants in the
present study, who said that “translators […] are seen as mere machines. Clients just think that with a click
of the mouse the translation is done […]”. Maybe we could use the results [of this study] to make them
more aware of our ‘hard work’. This sentiment is echoed in Karamanis et al.’s study (2011:10) where, in an
LSP context, in-house translators admit that “even though the client may receive a few queries, the full
scope of problems faced by the translators is not revealed”, thus the client is not aware of the amount of
work that has to be done to resolve all translation problems.

Therefore, in view of the amount of time translators spend on research activities and the lack of
comprehensive solutions designed to streamline these activities, a solution to reducing the time spent on
research would clearly be welcomed by the community of professional translators.

8.2. The use of resources

Although the use of resources by translators has been the object of several studies, the present study is the
first one to study the use of resources using two different samples and, in the case of the smaller sample,
to consider the self-declared and the observed data. The findings related to the use of resources are
associated with RQ2 and contribute to a better understanding of the volume and diversity of resources.

---

229 It could be argued that their research time could be lowered by using customised MT; however, such technology is
not generally available for freelance translators.

230 The concept of ‘shadow work’ has been recently discussed by Lambert (2015) in the context of doing unpaid work
once done by paid professionals, such as scanning our own groceries, filling our cars with petrol or booking a
holiday.
used by professional translators, pointing to the need to integrate these resources into translation environments.

As mentioned throughout the present work, the importance of research activities during the translation process cannot be underestimated. Even though experienced professional translators might rely more on context than on external resources to produce target texts, they still need to check if their hypotheses are correct, especially in a new context. For example, one of the terms in the translation task of the Main Study was ‘haircut’. The audio recordings revealed that the majority of the translators did not know the exact definition of a ‘haircut’ in the financial context and yet only eight translators researched the term, with the rest relying on the context and using TL solutions that referred to cutting, pruning or even confiscating. Whilst the use of a more generic word did not change the meaning, it could be argued that the full and exact meaning of the term ‘haircut’ as ‘a reduction in the stated value of an asset’ was not adequately conveyed in the target text in those cases. Therefore it is clear, especially for freelance professional translators who do not always receive reliable TMs or termbases from their clients, that external support is of great importance and it is highly unlikely that the need to use external resources during the translation process will ever be eliminated.

The results of the present study confirm that professional translators are now using predominantly online resources, with a small percentage declaring that they use mostly paper and other resources such as CD-ROMs. This, however, could depend on the availability of resources in a particular language. Zheng, for example, reports that in his study of novice, semi-professional and professional translators, most of the consultations (almost 76%) were carried out in what Zhang calls a Software dictionary and only 3% of all consultations belonged to the ‘other’ category, which included the Internet (Zheng, 2014:11). This implies that although online resources are dominant in many languages, in some languages electronic dictionaries still service much of the research process. As noted in section 5.2.2., three participants in the Main Study used off-line electronic dictionaries and one used a paper dictionary.

Regardless of the medium, translators use a great variety of resources, both linguistic (i.e. those specifically designed for language users such as dictionaries, glossaries or term banks) and extra-linguistic (i.e. those which are designed for general public use such as encyclopaedias, web pages or search engines). A total of seventeen different types of resources were observed to be used in the Main Study, including: monolingual dictionaries, bilingual dictionaries, synonym dictionaries, collocation dictionaries, orthographic dictionaries, idiom dictionaries, glossaries, term banks, thesauri, concordancers, discussion fora, non-customised machine translation engines, search engines, knowledge-based resources, web pages (parallel texts), online documents (pdf, ppt, digital documents) and ask an expert. Individual translators used between 0 and 12 resources, with a small percentage declaring that they use mostly paper and other resources such as CD-ROMs. This, however, could depend on the availability of resources in a particular language. Zheng, for example, reports that in his study of novice, semi-professional and professional translators, most of the consultations (almost 76%) were carried out in what Zhang calls a Software dictionary and only 3% of all consultations belonged to the ‘other’ category, which included the Internet (Zheng, 2014:11). This implies that although online resources are dominant in many languages, in some languages electronic dictionaries still service much of the research process. As noted in section 5.2.2., three participants in the Main Study used off-line electronic dictionaries and one used a paper dictionary.

Regardless of the medium, translators use a great variety of resources, both linguistic (i.e. those specifically designed for language users such as dictionaries, glossaries or term banks) and extra-linguistic (i.e. those which are designed for general public use such as encyclopaedias, web pages or search engines). A total of seventeen different types of resources were observed to be used in the Main Study, including: monolingual dictionaries, bilingual dictionaries, synonym dictionaries, collocation dictionaries, orthographic dictionaries, idiom dictionaries, glossaries, term banks, thesauri, concordancers, discussion fora, non-customised machine translation engines, search engines, knowledge-based resources, web pages (parallel texts), online documents (pdf, ppt, digital documents) and ask an expert. Individual translators used between 0 and 12 resources, with a small percentage declaring that they use mostly paper and other resources such as CD-ROMs. This, however, could depend on the availability of resources in a particular language. Zheng, for example, reports that in his study of novice, semi-professional and professional translators, most of the consultations (almost 76%) were carried out in what Zhang calls a Software dictionary and only 3% of all consultations belonged to the ‘other’ category, which included the Internet (Zheng, 2014:11). This implies that although online resources are dominant in many languages, in some languages electronic dictionaries still service much of the research process. As noted in section 5.2.2., three participants in the Main Study used off-line electronic dictionaries and one used a paper dictionary.

231 It appeared in the following sentence: This was evident in April when it is believed that a number of large deposit holders in Cypriot banks exchanged their money into bitcoins in order to avoid the ‘haircut’ enforced upon them by the European Union.

232 At the time of the Main Study data collection, i.e. between July 2013 and March 2014, the definition of ‘haircut’ in the financial context was not readily available. In response to the query define: haircut, Google search only returned a definition related to hairdressing. This has now changed and a second definition ‘a reduction in the stated value of an asset’ is being displayed by Google.

233 In a private communication related to this research, it was clarified that the Software Dictionary refers to the Kingsoft PowerWord CD-ROM dictionary, widely used among Chinese student/professional translators. There is an online version of the dictionary now, but it is reportedly “not as helpful as the installed program”.

245
types of resources, with an average of 6.25. This supports the accepted view that professional translators use “diversified consultation methods in translating” (Zheng, 2014:13).

Out of these resources, dictionaries were still very much used and, in line with other studies (e.g. Désilets et al., 2009) the use of bilingual dictionaries was observed to be heavier than the use of monolingual ones. However, text-based resources were found to be more dominant as a type in the present study (see section 8.4. for a more detailed discussion). No purpose built corpora were seen to be used in the present study despite the fact that almost half of the participants in the Main Study self-declared that they use them frequently or mostly (see section 5.2.3.). Surprisingly, the reported use of machine translation was higher than expected, given the time of data collection (2012/2013). The Global Survey of 540 translation professionals revealed the self-declared use of non-customised MT such as Google Translate by 53% of the respondents, whilst 12% said they use customised MT solutions. In the Main Study, much larger proportion of the sample of sixteen translators declared using non-customised MT (75%) and, similarly to the Global Survey, 25% declared using customised MT solutions234. The observed use of MT in the Main Study was lower than the declared as 25% of the translators actually used MT in one form or another during the translation task.

With regard to the use of specific resources, the freelance professional translators in present study, like the in-house professional translators in Désilets et al.’s (2009) study, had an “astounding” variety of resources in their tool box (Désilets et al., 2009:np), on average 10.2 per translator per task, almost identical to the figures in Désilets et al.’s study. The most used specific common resources, excluding websites specific to the subject of the translation task, were: Google search engine, Wikipedia, Linguee, ProZ, WordReference, ProZ, Google Translate and IATE.

Interestingly, the number of resources that the respondents had never heard of was high, both in the Global Survey and in the Main Study (see Figure 5.22 and Figure 5.23 in section 5.2.4.). This is despite the fact that many of these are specifically aimed at translators. This might suggest that many resources for translators are underused, perhaps because of a lack of adequate marketing or insufficient coverage during translator training, a point also made by Massey & Ehrensberger-Dow (2011b:20). Another possible explanation for the large number of resources that remain unknown to translators could be that some translators are indeed happy with just a few resources in their tool box and therefore do not venture out to find new ones. In the Main Study, five out of sixteen translators had five or fewer specific resources in their tool box.

It is possible that the market is close to being saturated with good quality resources in many language pairs and that many translators are content with what is available. Indeed, one of the questions asked in the profile questionnaire of the Main Study related to the level of satisfaction with the amount and quality of resources available in the translator’s own language pair. Almost 65% of the respondents were either satisfied or extremely satisfied, 24% were somewhat satisfied and only 12% were not very satisfied. For example, T1 reports in the profile questionnaire that “there are hundreds of good sources for Spanish, so I think it is well covered”. None of the respondents expressed complete dissatisfaction with the resources on

234 The use of customised MT solutions should be taken in the context of the make-up of the sample in the Global Survey which included a variety of translation-related professions such as translators, interpreters, terminologists, project managers etc. and some of these professions could have been exposed to customised MT solutions.
offer in their language pair, although a handful of participants (translating into Polish, Hungarian and Dutch) were ‘not very satisfied’.

The relative satisfaction with existing resources would probably explain why translators are not willing to support the development of new technologies specifically aimed at them. Zetzsche has recently commented on a Linguee development that attempted to “really reach out to the professional community by offering a f-a-n-t-a-s-t-i-c tool that searched through the corpus and took into account all the context of whatever you searched for right from your translation environment tool, thereby producing vastly superior results” (Zetzsche, 2015a, original emphasis). Zetzsche bemoans the fact that the community of professional translators “effectively rejected that offer because of its price (something like €5 a month), so it was a natural move for Linguee to move on and focus on a general audience” (ibid.).

However, whilst professional translators might not be quick to support new developments, they certainly feel the need for the resources to be integrated into their environments (Zetzsche, 2015b). Given that translators spend so much time on research activities and that they use large and diverse quantities of online resources, it is surprising that still very little integration of these resources can be found within translation environments. Other researchers have also expressed the need for such integration. Zapata points to the “need for a tighter integration of Internet-based informational tools and translation environments” (Zapata, 2015:151) which seems to echo Désilets et al.’s call for “complementarity and synergy [...] [and] unified interfaces” within the new translation technologies (Désilets et al., 2009). Therefore, as noted earlier, a solution aiming at the reduction of research time and integration of the online resources into the translation environment is needed.

8.3. Research distribution

The features relating to research distribution in the translation process such as non-linearity, iterativeness, overlapping and parallel processing have been discussed in TPR literature (e.g. Séguinot, 2000; Lauffer, 2002). However, they have not been discussed strictly in relation to research activities which are considered to be a separate type of activity (i.e. information-seeking) and not from the point of view of information behaviour, a research gap the present study attempts to address as part of RQ3.

The findings of the present study demonstrate that the distribution of research within the translation task is a complex issue. Translators were observed to flick backwards and forwards between the text and the resource (backtracking), conducting research retrospectively (long after they encountered the problem) or prospectively (even before they encountered it). They also researched indirectly, i.e. looked for potential candidate terms whilst reading parallel texts. The data shows that research episodes can happen at any time during the translation process and that they can be clustered during the different phases of the translation process (beginning, middle or end) or can be evenly distributed throughout the entire process. Research episodes can be very simple (a one-step consultation) or composite (multi-step consultations in several resources). Furthermore, a translation problem might require a translator to revisit it at a later stage during the process, sometimes more than once, resulting in the formation of research sessions, which can be composed of several research episodes relating to the same problem or research unit (RU).

235 T6 expressed satisfaction with EN-NL resources, but was dissatisfied with the FR/NL and the DE/NL resources, especially bilingual dictionaries (too old or lack of).
Such variations are ignored in some latest innovative cloud-based solutions (e.g. SmartCAT\textsuperscript{236}), where a reviser can work on a translation as soon as the translator has finished working on a segment, thus eliminating the possibility of carrying out any retrospective research or end revision.

A detailed quantitative analysis of the observed data in the screen recording part of the Main Study revealed that the research steps were mostly very short, with an average consultation lasting 20'', and the range being between 12'' and 38''. The length of research episodes averaged 1'1''\textsuperscript{237} and only a few research sessions were composed of multiple research episodes. Because of the complexity introduced by the presence of the research sessions when carrying out the analysis of data, many calculations were actually performed in relation to the individual problems, i.e. in relation to research units (RUs). The data has shown that on an average 2.6 research steps are taken to resolve a problem related to a RU by consulting 1.8 specific resources\textsuperscript{238}.

Considering that research episodes which on average lasted 1'1'', were interlaced with translation episodes which on average lasted 1'27'', the translation process could be perceived as highly fragmented, with constant switching between translation and research activities. These switching patterns can differ considerably between translators and the switching could be less frequent if a translator is very familiar with the domain and judges that not much research is needed (as in the case of T2 who did not conduct any research). However, in most cases, to the observer, the translation process could still be perceived as a rather non-linear, iterative and in some cases relatively meandering flurry of activities which translators manage to integrate into a seamless task. Therefore, coping mechanisms must exist to manage the disruptive potential of the interactions with resources. In the present work it is argued that professional translators cope by developing different ‘styles’ which are compatible with their natural information behaviour traits. For example, some translators plan their research, whilst others simply ‘absorb’ the information they encounter through meandering behaviour. Some might manage by fast and shallow processing and others by simply avoiding any research that is not deemed absolutely necessary thus reducing their research activities to a minimum. These different ways of approaching and managing research activities are the backbone of the Taxonomy of Translator Research Style discussed in detail in Chapter 7 and summarised in section 8.5 below.

Although the present research is not concerned with cognitive activities, such perceived ‘disruptiveness’ will no doubt have repercussions for the cognitive load involved in the switching between translation activities seen as problem solving and decision making, and research activities which are governed by information-seeking behaviour. Clearly, research is needed in this area, especially now that translation activities are acquiring a new dimension associated with post-editing (modification of pre-translated text rather than generating own text) and research activities change to accommodate cognitive processes related more to ‘selecting’ as opposed to ‘generating’ solutions (Austermühl, 2013; Pym, 2013).

\textsuperscript{236} https://www.smartcat.ai/
\textsuperscript{237} Comparable figures are reported by Enríquez Raido (2011) and Volanen (2015).
\textsuperscript{238} The number of steps within a research episode does not equal the number of specific resources used in that episode as a translator might for example reformulate a query using the same resource in which case the resource is counted once, but two steps are calculated.
8.4. Variations in resource behaviour

The main theme emerging from the present study is the effect individual differences might have on the research activities within the translation process and the attendant hypothesis that the differences in translator performance might reside not just in the level of expertise or domain knowledge, but also in other aspects such as personal preferences and natural predispositions. This aspect of translator behaviour during the translation process has not been studied before and, as a result of investigations prompted by RQ3, the present study offers a significant contribution to the understanding of translator online information behaviour. One area of this contribution relates to the resource behaviour, the other to the research behaviour of freelance professional translators. In this section, the conclusions and discussion relating to the former will be considered, based on data from the Main Study.

The preferences and natural predispositions of professional translators with regard to resource behaviour have to be considered in the context of the developing technologies and the translators’ adaptation to the changing nature of available resources. It could be argued that because in the pre-Internet era the choice was much narrower and constrained by the limited availability of certain paper resources, translators’ resource behaviour would have been much more homogeneous. The increased number and variety of online resources and their greater availability and accessibility in the online medium has resulted in a greater choice of options for translators and thus in a more diverse resource behaviour.

To the best of my knowledge, resource behaviour has not been systematically investigated before in an empirical fashion. Fulford & Zafra (2004) analysed the adoption of Internet services by freelance translators; however, their account is based on a survey rather than on observation. They identified three ‘clusters’, based on the frequency of use of certain technologies: data and information repositories, information retrieval and processing programmes and communication tools. These clusters were described rather vaguely using qualifying adjectives such as ‘extensive use of …’ or ‘some awareness of…’ and the data was analysed against variables such as age, experience and education and did not examine the clusters in the light of individual differences.

The present study’s Resource Type User Taxonomy (RTUT) is an attempt to capture the diversity of translator resource behaviour (resource styles) and to then examine whether variables such as age, experience, gender or familiarity with the domain can be associated in any way with these styles or their individual components. The RTUT captures the use of resources according to their nature (i.e. termino-lexicographic vs. text-based), the research direction (i.e. ST-oriented, mostly associated with the use of termino-lexicographic resources vs. TT-oriented, mostly associated with text-based resources239) and the way these resources are accessed (i.e. top down via keyword in a search engine vs. bottom-up via a known resource). The combined effect of these three categories resulted in the classification of the professional translators in the Main Study into Dictionary Enthusiasts (six participants), Parallel Text Fans (four participants) and Mixed Users (five participants). Four participants were also classified into a separate MT User category based on their use of this resource during the task. Although translators seem to be relatively evenly spread across the different RTUT types based on all three RTUT components, some results from the individual components are worth pointing out.

239 Although many ST-oriented searches were also conducted in text-based resources.
The first important finding is that in the present study translators used more text-based resources than termino-lexicographic ones. As already noted, Désilets et al. (2009) report on the increased use of text-based resources which “have made it to the mainstream [...] [but] have not replaced the termino-lexicographic resources”. However, the main difference between the two studies is that text-based resources were used more extensively than termino-lexicographic ones in the present study (twice as many in numbers and twice as frequently\(^{240}\), perhaps because Désilets et al.’s study was carried out five years before the present study. A sharp increase in the use of online parallel texts was also observed by Hirici (2013) in her longitudinal study of resource use by trainee translators.

It was also noted that in the present study the younger translators, who unsurprisingly had less experience, used more text-based resources than their older and more experienced colleagues. Therefore, it could perhaps be argued that the new generation of translators who were trained and developed their careers with the Internet as a main source of reference aids are showing more text-based dominance. Also, familiarity with the domain was revealed to play some role in the choice of resources as slightly more text-based resources were used by those who had some experience in the business/economic/financial domains, implying that they might have been more interested in research activities to gain more background knowledge or to check the usage of particular terms rather than in equivalence seeking. Surprisingly, male participants were found to choose termino-lexicographic resources more often than female participants, but the explanation for this goes beyond the scope of the present thesis (see section 7.3. for more details on all the above mentioned variations).

The dominance of a ST-oriented research direction within the sample of experienced professional translators could also be considered noteworthy as it is generally assumed that with greater experience translators acquire the ability to detach themselves from the source text thus being able to formulate hypotheses about solutions in their target language more easily. However, the present study shows that more ST-oriented searches were generally conducted during the translation task in the Main Study and no pattern was found associating the age or experience of the translators with the directionality of their searches. The only exception was gender as it was found that more male than female participants displayed ST-oriented dominance (75% of men\(^{241}\) vs. 45% of women). As discussed in section 6.1.2., the findings relating to the directionality of research are somewhat surprising and contradict findings from Enríquez Raído (2014:84), Prassl (2010) and Volanen (2015). Furthermore, lack of familiarity with the domain and hence with the terminology, did not prevent the two most prolific TT-oriented searchers (T11 and T15) from formulating queries in their target language. Therefore, it could be argued that the directionality of research does not have to depend on experience or familiarity with the source-text domain, but could also be part of one’s resource style. As it is generally believed that TT-oriented research fosters more fluent translations, further exploration of the relationship between research direction and the quality of the translation product is needed.

Regarding the third component of the RTUT, research strategy, the findings suggest that a bottom-up search strategy, i.e. starting the search from a known resource rather than a keyword in a search engine is more prevalent, with almost half of the Main Study sample displaying bottom-up dominance and the rest displaying a mixed or top-down dominance. As in the case of research direction, no relationship between

\(^{240}\) These differences could, of course, be down to the way the use of resources was measured or the way the resources were categorised.

\(^{241}\) T2, who was also a man, could not be accounted for as he did not conduct any research.
age or experience and research strategy was found, and gender was the only variable that seemed to have an influence on the research strategy, with 75% of male participants choosing to start the search with the known resource. However, a number of conclusions can be drawn from the observed bottom-up dominance with regard to the way the search is initiated. First of all, the fact that translators start with a known resource indicates that they know what resources are suitable to satisfy a particular research need. With a few exceptions where extra-linguistic knowledge was sought in dictionaries, all translators were comfortable with resources they used and knew well. Echoing Désilets et al., translators “seemed to know very well the strengths and weaknesses of the various resources in their toolboxes” (Désilets et al., 2009:np). T1 reports in the profile questionnaire that she “researched thoroughly through trial and error and compiled a good list bookmarked for easy access”. A similar account is given by T12, who says that “these resources were practically non-existent during my first years as a translator, so I mostly trained myself by trial and error [and] still regard them as a blessing”. T15 confirms that “there are many tools available in my language pair and it is quite easy to triangulate meaning for this reason. Their quality is sometimes doubtful, but it is easy to discriminate between doubtful and reliable sources in general in my language pair”. However, some translators felt that they would benefit from updating their knowledge on the available resources. For example, T13 and T7 explicitly said that the profile questionnaire had inspired them to explore more resources and T16 felt a “bit behind”. Despite that, the overall feeling was that professional translators have their own preferences as to which resources they like to use and know what to expect from them. Many comments related to Wikipedia, the second most used resource in the present study. Translators were aware of the ‘community’ status of Wikipedia but considered it a trustworthy resource. For example, T12 mentions that despite the fact that academia frowns upon Wikipedia, he finds it very useful as a source of general information. While researching New York’s local currency ‘Ithaca Hours’, he comments on the fact that Wikipedia has the same information as the original source, i.e. ‘Ithaca Hours’ website and much more.

The finding that more ST-oriented dominance was observed alongside more bottom-up oriented behaviour could be seen as logical as no ‘reverse’ use of dictionaries, i.e. checking or confirming a TL hypothesis in a dictionary or a glossary to seek equivalence was observed. However, the fact that more text-based dominance accompanied the ST-oriented/bottom-up behaviour, though it is associated with more target-oriented research activities, could be seen as somewhat confusing. This could be related to methodological factors. For example, 60% of the one-step research episodes (21% of all research episodes) were one-step consultations in dictionaries or glossaries. This would mean that 21% of all searches were assigned a 1-1-1 correspondence, i.e. one ST resource/ one ST-oriented search/one bottom-up strategy, whereas in the many multi-step consultations starting with a search engine, where no additional keywords were used due to the translator clicking through multiple pages without reformulating the query, the text-based resource was still counted as one, but there were no corresponding counts for research direction or research strategy. Therefore, the count for text-based resources is higher in proportion to the counts made in relation to research strategy or research direction, and, consequently, higher than the count made for termino-lexicographic resources. Furthermore, this discrepancy could also be due to the way resources

---

242 However, due to time constraints no systematic calculations were carried out to cross-examine this discrepancy.
were calculated as a type\textsuperscript{243}. Therefore, the methodology for calculating the use of resources with regard to their nature needs to be refined in future studies to better represent the distribution of resources against research strategy and research direction.

Although currently the bottom-up approach still prevails amongst professional translators, the one third of searches initiated via search engines could suggest that the use of top-down techniques might increase in future at the expense of bottom-up techniques. At the moment search engine capabilities are not sophisticated enough for use by professional translators whose research needs are very different from those of general users or other professions. However, certain developments could pave the way for search engines to become much more useful for professional translators. One such development, for example, is based on a closer integration of translation technologies and online resources with web searching, i.e. the embedding of MT solutions (MacKenzie, 2015:online) and lexical resources such as dictionaries (Rundell, 2013:7) into search engines. Lexicographers now talk about the notion of ‘content strategy’ which will transform the landscape of resources, including the possible disappearance of dictionaries as stand-alone resources. Some efforts have already been directed at creating a new type of content in the form of a ‘lexical hub’ that can be intelligently read by machines through semantic annotation and can adjust to the current technologies and user needs (Pearsall, 2013:4). As Pearsall, an editorial director at OED, convincingly argues,

dictionary publishers are in the business of fulfilling the need, that of providing resources to help users to understand or use language. It just so happens that dictionaries have been a good vehicle in the past to fulfil that need. But this may not be true for the future, or for the present. If we focus on what is needed, who needs it and why it is needed, we might well redefine what we are doing and even dispense with the idea of the Dictionary itself along the way.

(Pearsall, 2013:3)

Disruptive as it may seem, moving away from traditional resources such as dictionaries and embracing new technologies could accelerate ‘discoverability’ and support translators with new, faster and more efficient ways of satisfying their research needs. For now, it seems that the tighter integration of online resources and search engines has not yet materialised and search engines are not sophisticated or personalised enough to cater for all translators’ needs. The way Google works at the moment is not ideal for translators as it prioritises websites according to the number and the quality of internal links, i.e. the ‘connectivity’ to other websites. A ‘less connected’ website with relevant results for translators has little chance of appearing at the top of the search results page. This is compounded by the fact that translators rarely look beyond the second page of the results, unless they display Explorative or Prolific tendencies. A solution with more potential for translators is currently in development by IBM. The new search engine ‘Clever’ works in a similar way to Google in the first phase of the search, but in the second phase “unlike Google which retains rankings for individual websites in its index, independently of the user’s query, Clever will always […] prioritise each page according to the context of the specific search statement” (Green,

\textsuperscript{243} For example, if a dictionary was used, it was counted as one type on a par with any single Web page consulted. Whilst translators would normally use one or two dictionaries but use them regularly, they might use multiple Websites but only on one occasion. To balance this and to arrive at a more accurate portrayal of the use of resources according to their nature, frequency counts were used (see Table 6.16). Nevertheless, this could still have some kind of effect on the typology.
It means that once a website with relevant content has been accessed, it will boost the search ranking for the particular keyword independently of its linkage structure.

Apart from future developments in search mechanisms, many changes driven by the developments in the semantic web, information retrieval, artificial intelligence and related areas will be shaping the future ‘resource type user’ and we can predict that in future all the contents of linguistic resources will be much more easily accessible through search engines. It is therefore in the interests of translation process research and the future development of translation tools and resources to monitor patterns of interaction with online resources by professional translators and to see, as Zapata (2015:151) suggests, “how the interaction with [...] information tools can change over time”. A longitudinal study observing the use of search strategies over the five to ten years would bring much insight into this area, with far-reaching implications not only for the development of resources but also for training translators in the area of search optimisation for translation.

8.5. Variations in research behaviour

Like the previous section, this section is concerned with the findings that relate to the influence of individual differences on the research activities within the translation process. Here, a summary of the findings relating to the research behaviour of the study participants will be considered. These findings form part of the answer to RQ3 and are based on the Main Study data.

The main contribution to translation process knowledge resulting from the analysis of translators’ research activities lies in the recognition that the variations in translators’ research behaviour may not be entirely dependent on variables such as expertise or domain knowledge but may also be influenced by personal preferences and natural predispositions. Furthermore, it is also argued that whilst resource behaviour is dependent on technology, research behaviour has more to do with information behaviour, seen as an innate property of human nature.

As noted in sections 5.2.4. and 7.1.1.3., individual differences in translator research behaviour can already be seen in data obtained by Nord in 1997 in a study of the use of resources by professional translators in the pre-Internet era (Nord, 2002). For example, the number of resources used by individual translators varied between 1 and 7, the number of individual consultations (steps) per hour varied from 3.2 to 40.8 and the average time between these consultations varied between 1’28’’ and 18’45’’. This shows that, indeed, individual differences between the use of resources by translators, whether paper or online, are hard to ignore and were observable even in the pre-Internet era. However, although these differences can be seen in Nord’s data, she did not explicitly comment on them, possibly because the texts being translated in her study varied enormously (from power cuts in the US through Japanese customs to a request for structural changes to a building). In a more recent study, Enríquez Raído succeeded in identifying two web searching styles, the ‘shallow’ style associated with “horizontal, checking and comparing” behaviour and the ‘interactionistic’ style where searches are found to be deeper and wider (2014:139-140). The present study also recognises these characteristics, however, they are not found to be entirely attributable to length of experience, but also to translators’ individual ways of approaching research activities.

Building on the above findings, the present study presents an attempt to provide a systematic approach to analysing translators’ research behaviour during the translation process. The analysis of translator research behaviour was carried out by considering what I have called ‘primary’ and ‘secondary’ categories. We can
recall that the primary categories, calculated for all participants in the Main Study relate to the volume of research and resources used, the diversity of resources, the intensity of research (measured in the number of steps per research unit), the pace of research (measured in relation to the length of a research step and research episode) and the time spent on research. The secondary attributes communicate aspects of research behaviour such as depth of research, shape of research paths, planning and many other characteristics of research behaviour observed in more than three translators. The secondary categories were used in the analysis to either reinforce the research style assigned to a translator as a result of the primary categories, or to shift their position in borderline cases. The resulting Typology of Translator Research Styles (RTUT) places translators along two axes intersecting in a matrix: volume- and time-related features, thus creating five distinctive research styles as explained in section 7.1.1.7. These five styles, shown in Figure 7.6., were named as Economical, Understated, Methodical, Prolific and Explorer.

Although five research styles were identified, the participants were not evenly distributed within the matrix. The more ‘voluminous’ styles, i.e. Prolific (high research volume, high research speed) and Explorer (high research volume, low research speed) are the least represented, followed by Understated (low research volume, low research speed) and Methodical (moderate research volume, moderate research speed). The most represented category, the Economical Translator (low research volume, high research speed), covered half of the sample.

The fact that half of the sample congregated in one section of the TTRS grid of the matrix and displayed Economical research behaviour linked to low research volumes and low research timings is of some importance for understanding the way the translation process happens in the real world of professional translators (as opposed to trainee translators or students) as well as for translation pedagogy, i.e. what are the priorities in the classroom in the light of these findings? As no previous studies of this type are available, it is not possible to observe whether this clustering around the Economical Translator style is associated with the growing demands on translators in terms of ever growing volumes delivered in ever shorter times, where a compromise needs to be achieved between translators’ natural working patterns and market demands. As already mentioned in 7.2.4., one of the Economical Translators (T16) expressed such a view by saying that she feels the need to adapt to the way the industry works by dealing with the text in a faster and more efficient way, for example, by skipping the orientation phase. Interestingly, T16 was originally placed on the border with the Methodical style which is associated with more careful planning of research activities and was also close to the border with the Prolific translator style due to the volume-related aspects of her translation. This would corroborate the above mentioned hypothesis that market demands might interfere with translators’ natural predispositions and working styles. This hypothesis could be linked to research into the so-called Google generation (Rowlands et al., 2008) where shallow, horizontal, ‘flicking’ search behaviour has been attributed not only to younger generations but to all web users. As Rowlands et al. conclude, “all of us have changed the way they (sic) seek information. We are all Google generation, the young and the old, the professor and the student and the teacher and the child” (Rowlands et al., 2008:308). Furthermore, they argue that “a fundamental shift in the way people seek and read information has already occurred and that the impact of the shift has yet to be understood by information providers and educators” (ibid.). Perhaps the clustering of the professional translators in the Economical section of the TTRS grid supports their observations.
The second most densely populated section in the TTRS grid is the moderate Methodical Translator who needs a degree of structure and planning in their research. The need for planning as a risk management strategy has been proposed by Gouadec (2002) who supports the extensive planning approach to translation by means of job descriptions and pre-translation analysis, designed to reduce (translation and research) effort as a risk management strategy. Pym (2010:np) disagrees with this excessively systematic approach on the basis of the unnecessarily long procedures and cost in relation to risk. He convincingly argues that there “must still be room for experience, pragmatism, justified non-translation, creativity and inspiration” (Pym, 2010:np). The results of the present research would suggest that both Gouadec and Pym have a point in that how translators manage their translation process might depend on their translation style as well as on their research style. Therefore, any discussion around risk reducing strategies and management of the translation process also need to take into account that translators are not a homogeneous group, but individuals with their own preferences and styles.

Thus, it could be argued that Methodical Translators will thrive in an environment that requires some planning, for example, in technical or medical areas where terminology work is more strategic, but that Explorers will struggle to come up with a plan as they work in a discovery, rather than planning mode and would therefore be suited to less time-constrained work, for example in literary translation or in transcreation. Similarly, while the Explorer is likely to make a poor MT post-editor, Economical translators will be more suited to this kind of work, being efficiency-driven and having no desire to wander off the chosen path. Having said that, it is also quite possible that most translators can adapt to the demands of a particular job and move along the volume- and time-related spectrum according to the individual circumstances of the assignment. Such skills need to be taken into account in translator training. Nevertheless, the idea that professional translators can ‘attune’ their performance according to market demands should not obfuscate the fact that, in ‘ideal’ conditions, translators are likely to display their natural styles in accordance with their inherent predispositions because information behaviour, as Spink & Heinström (2011) suggest, is an innate feature of human behaviour. As such, all types of research behaviours need to be valued and encouraged to thrive, unless we resign ourselves to becoming more machine-like in our working patterns, succumbing to the pressures and demands of the market. If the human way of doing things is to be preserved, all our efforts need to be put into creating technologies that support human behaviour rather than coercing humans to adapt to technology without considering the cognitive and ergonomic impact it might have on them. As Heinström insightfully puts it,

No matter how fast information technology evolves or how sophisticated search systems we learn to master, our basic human reactions remain as they have been through centuries. Our behavior, even in seemingly rational activities such as information seeking, is influenced by our holistic being as a creature of physiological, cognitive, and affective processes. This characteristic has important implications for the development of user-centered information services. The user can learn to adapt to search systems, but more importantly search systems should be adapted to users’ natural ways of seeking information. The traditional approach in library and information science has been to support users to overcome their possible weaknesses in search ability. Equally important is to recognize particular strengths in the users’ habitual ways of information seeking and adjust systems to support these tendencies.

(Heinström, 2006a:1440).

As already mentioned in previous sections, the need for a solution that would streamline research activities within the translation process through the integration of online resources and/or the provision of more sophisticated search facilities is tangible. But the results of the present study strongly suggest that such
developments must take translators’ resource and research behaviour into account by providing solutions adaptable to their particular working styles. As Zapata aptly concludes: “there is a tangible need to design and develop ergonomic and flexible interfaces that take the human factor into consideration and that are adapted to the translator’s workflow and needs” (2015:142).

8.6. Externalisation of the translation activities

Although cognitive processes per se were not examined in the present study, some cognition-related observations pertaining to RQ3 can be made. One of the main points that has surfaced relates to the influence that external resources have on the translation process and the associated cognitive ability (or choice) to retain information. As some early process studies have shown (see section 2.1.), the use of paper dictionaries was found to be ‘interrupting the flow of translation’ and some researchers even prohibited the use of resources in their studies “to prevent problem-solving processes from being blocked or broken off too early by simply adopting ready-made solutions” (Lörscher, 1991:40). Translation, it seems, was truly happening in the translator’s head and the professionals would consult reference material only in critical situations, as opposed to students who were found to overuse dictionaries. With the instant availability of online resources and ever increasing recall precision, translators might be more tempted to look up information online rather than in their own memory. Duval argues that certain cognitive functions such as memorising can be affected by intensive use of the Internet as people “will turn to the Internet to ask even the simplest question. Just knowing that a piece of information is readily available anywhere and anytime leads humans not to memorise” (Duval, 2012, in: Zapata, 2015:152). In the context of translation, Pym refers to this phenomenon as ‘extending’ and ‘externalising’ human memory which is believed to be one of the dimensions of the translation process most affected by technology (Pym, 2011:1). T6 explicitly confirmed this in the post-task questionnaire, saying that “incidentally, since I use Déjà Vu, I found that I am much less motivated to even try to remember certain terms translations of when and where I used them, knowing that Déjà Vu will bring them up when needed”. Although Pym and T6 refer to TM, this observation can be easily extended to the use of online resources as a quick ‘aide-memoire’, a convenient and fast alternative to retrieving options from one’s own memory. This kind of behaviour was observed in T3 who looked up a high proportion of general words such as ‘protect’, ‘visible’, ‘evident’, ‘freedom’ or ‘potential’, all in bilingual dictionary consultations that lasted two to three seconds, despite rating her retentive memory as high (four out of five, five being the highest). Furthermore, the results of the present study, compared to the study carried out by Nord in the pre-Internet era (Nord, 2002), show that the translation process observed today is interrupted by a consultation more than twice as often, i.e. every 1.5 minutes now compared to every 3.5 minutes back in 1997 (see section 5.3.1.), which would suggest that translators rely more on online resources when they are more readily available.

This externalisation and extension of translators’ memory through technology entails changes in cognitive processing during the translation process, from a situation where translators had just a source text in front of them to one where many windows displaying the source text, the target text, the concordance, the termbase, the glossary, the MT suggestions, the predictive text etc. are competing for their attention, i.e. a situation where they are literally bombarded with ready-made solutions at any one time. Arguably, translation now requires the deployment of different cognitive processes to manage the changing ergonomic and techno-cognitive dimensions of the translation.
Pym (2013:493) maintains that a “profound shift […] has been occurring progressively with the impact of the Internet”, namely from identifying possible solutions to translation problems using generative cognitive processes on the one hand to selecting between available solutions using selective cognitive processes on the other hand (ibid. original emphasis). It is possible that the amount of translation that actually happens in the translator’s head will decrease with time to be replaced by content leveraged through various technologies such as TM, MT or external resources. The repercussions for human creativity, individuality and the general change in language caused by the use of more and more repetitive language patterns that are re-used through the use of technology cannot be underestimated. These issues remain to be tackled in future research.

8.7. Practical implications of the findings

Many of the implications of the present study were touched upon in the previous sections whilst discussing relevant findings. Here, the more practical implications of the findings of the present research will be discussed.

8.7.1. Technology development

The main implication of this study is that online resources are in need of integration and research activities within the translation process in need of streamlining. Although the integration of certain resources is already happening (see for example BabelNet in section 4.3.) and search tools to mine resources are being developed (see for example IntelliWebSearch), a much more comprehensive approach needs to be taken in order to cater for translators’ research needs. The integration needs to go deeper not only horizontally across the many resources (Bowker & Marshman, 2009), but also vertically - into the individual CAT Tools244. Such integration would require a coordinated effort of collaboration and interoperability from many industry stakeholders, but the integration itself will not be enough. A solution, a kind of ‘intelligent, personal research assistant’ (IPRA) as I have called it, is required to mine those integrated resources in the way a human translator would, i.e. according to their natural resource and research style, but in a faster and more efficient way. The only way to achieve this is by grounding the IPRA in artificial intelligence (AI)245. Only a breakthrough such as this, based on embracing the decision-making processes of humans but deeply embedded in state-of-the-art machine learning, has the potential to make a difference in reducing the time spent on research activities. By doing so, as already noted, IPRA could offer the potential to streamline research activities within the translation process according to each translator’s individual preferences and style as each user could develop their own, unique interactions with the system. This could provide a truly translator-oriented technology geared towards time-saving, but one that also respects natural, human behaviour.

However, further research is needed in order to gain a better understanding of different research styles and the general online behaviour of professional translators. Bawden & Robinson also refer to this particular need for obtaining hard evidence on individual information styles. This, they argue, will allow

---

244 An example of the horizontal integration of an independent platform into various TEnTs is Slate Desktop, a personalised MT engine for translators, which is currently being developed as an on-demand, plug-in to various CAT Tools.

245 ‘True’ artificial intelligence (AI), i.e. one where the system is ‘conscious’ is not required for this type of assistance. A ‘weak’ type of AI based on machine learning would be sufficient to serve this purpose.
“the tailored provision of information, in ways most acceptable to the individual, with greater precision than current approaches, which are largely based on assumptions of the needs of particular professional or educational groups” (Bawden & Robinson, 2011:1 emphasis mine). Furthermore, once developed for the purposes of professional translators, IPRA could be adapted to other professions and user types.

To summarise, the present research brings us a step closer to a future solution for reducing the time spent on research activities and for developing ethical technology that thrives on natural human behaviour and yet increases productivity. In order to preserve our innate, human traits, efforts need to be made to design machines which will work in a way that is more human-like.

8.7.2. Translator training and teaching

The findings of the present research are also applicable to translator training and teaching. First of all, based on the findings related to the use of resources, we can focus on developing the competencies that rely on technology, for example, instrumental competence (PACTE, 2011), information competence (EMT Expert Group, 2009), technological competence (EMT Expert Group, 2009) and thematic competence (EMT Expert Group, 2009). As mentioned in section 2.3., drawing on the actual practices of working translators is the key to a successful transfer of ideas and the implementation of procedures from expert to novice. With the consent of participants, short videos could be made available for students to expand their declarative knowledge of the availability of various resources in their language pairs and their suitability for solving different kinds of translation problems, as well as their procedural knowledge as to how other translators search for information online. An important observation by Massey & Ehrensberger-Dow pertinent to this discussion is that “resource use and search skills [...] developed ‘on the job’ by many older practising professionals [...] may not always be either efficient or effective” (Massey & Ehrensberger-Dow, 2010:138). Therefore, the training provision for continuing education as well as for initial translator training is vital in the ever faster-paced technological translation environment of today.

The providers of translator education and training should be encouraged to seek inspiration in empirical research studying translator behaviour, as training provision is too often based on the experience of an individual trainer. The skills taught on university courses particularly need to reflect the skills required outside the university walls. As one of the participants said, she could not afford to implement some of the ideas she was taught on her translation course, for example, reading the ST thoroughly before translating it because of the tight time constraints in which she found herself working. Therefore, instead of teaching students to thoroughly examine the text before translating, they could, firstly, discover their natural behaviour during the orientation phase, i.e. whether they are planners, head starters etc., and secondly, for those who need to understand the text before translating it, skills such as power reading could be developed alongside translation skills and the use of text analytics software for translation purposes could be introduced into technology modules to facilitate speedy text comprehension.

Last but not least, the taxonomies that have emerged from the present study could be used in translator training as an awareness-raising exercise about students’ or indeed, professional translators’ natural predispositions, strengths and weaknesses by comparing their own performance with the performance of colleagues. For example, a strongly termino-lexical oriented translator could benefit from seeing how a text-based dominant translator explores parallel texts by finding potential solutions through prospective research and, conversely, a text-based dominant translator might try to use termino-lexical resources now and then to shorten the research time devoted to background reading. Similarly, a strictly Economical
Translator might learn from the Explorer that sometimes it pays to dig deep and, vice versa, the Explorer might learn that sometimes the quest for perfection may not result in the perfect solution. Often just seeing how other people proceed with the same task is revealing and potentially enriching for our own ingrained procedures. In fact, all participants in the Main Study were very interested to see the outcomes of this research purely for this reason. Massey & Ehrensberger-Dow also support such classroom activities. In their study, students who were asked to comment on each other’s screen recordings were particularly interested in observing and judging information behaviour. Their interview comments focused on the differences and similarities between their own and their peers’ procedures, and especially on improving efficiency, research techniques, resource use, workplace organisation, and source text analysis

(Massey & Ehrensberger-Dow, 2013:162)

Students commented on this activity as being a positive one noting that it “increased their ability to reflect on their own translation process” (ibid.).

Reflecting on a student’s own style, together with identifying other strengths and weaknesses, could also be of practical benefit in terms of choosing their specialisation or the right path within the translation industry after graduation. For example, the more explorative and creative types who trawl and meander through resources would very likely be well suited to literary translation or transcreation. They would probably not be so suited to post-editing as their threshold for perfection could be too high. The Economical translators, on the other hand, could possibly make much better post-editors due to their fast, shallow and undemanding research style. Choosing their specialisation by taking into account their own natural predispositions could help students when entering the profession after the completion of the translation course.

To sum up, mutual observation in the classroom combined with education about individual differences in research and translation styles could help students to better understand themselves and their working patterns and to relate to each other by appreciating individual differences. It could also help them to apprehend their own strengths and weaknesses which, in turn, would foster a more targeted skills development and thus facilitate their entry into the profession with the confidence that they possess the necessary skills to do the job. Last, but not least, the awareness of various working styles could also inform the work of project managers dealing with particular translators.

8.7.3. Research

As much as it is hoped that the present study contributes to a better understanding of the translation process by focussing attention on the importance of research activities during this process, it is further hoped that it will stimulate other researchers to engage with the findings and improve or develop alternative methodologies to test these findings on other, hopefully larger samples.

With regard to resource and research styles, the proposed taxonomies (RTUT and TTRS) can be considered to be the starting point to a further inquiry into these styles, informed by interdisciplinary insights from fields such as information behaviour, psychology and social sciences.
The importance of the present research lies in the opening up of many questions which can now be followed up using larger-scale studies focussing on particular issues raised by the present work, for example, whether research (and indeed translation) styles might exist independently of variables such as experience, familiarity with domain and, indeed, independently of the text type or task complexity.

In order to do this, two things need to happen. Firstly a bank of data containing genuine translation sessions from translators working in professional settings should be created and made freely available to researchers to download and upload data in approved formats. As mentioned in section 3.3.4., data collection was one of the most challenging tasks in the present study as a lot of time was invested in finding suitable candidates and dealing with the data collection. The first steps have already been taken in the form of the CRITT TPR-DB database (Carl, 2012a) which contains a wealth of process data, currently over 450 hours of text production time from 1400 translation, post-editing, revision, authoring and copying sessions in ten target languages. This database contains user activity data such as keystroke logging and in some cases eye-tracking recorded with Translog II and CASMACAT and some logs from the use of external resources collected through Inputlog. However, there is a need for a more comprehensive database containing other types of data such as screen recordings, video recordings, field notes, interviews or demographic data relating to the participants. Although much can be learned from quantitative data recorded by tools such as Translog or CASMACAT, these methods need to be complemented by data suitable for qualitative analysis. As Zapata correctly observes,

quantitative data […] may not be sufficient to investigate the TII [Translator-Information Interaction] process and evaluate the quality of information and information systems. Thus, it becomes necessary to further examine, through different data-collection methods such as input logging, screen recording, eye-tracking, active observations with video recordings, and interviews, different forms of behavior; and to formulate more detailed hypotheses about the TII processes and the usability of information tools.

(Zapata, 2015:150)

If more and varied process data could be made available for re-use, perhaps more researchers would be willing to design their studies around this data and try to replicate studies based on the same data but using a different methodology, or the other way round.

The second thing that needs to happen to further research into the translation process is the development of tools that could capture this process in translators’ natural environments. The lack of large-scale research could be down to the lack of convenient tools that could be used for large-scale observations of how the translation process unfolds, for example parallel with the use of online resources. As mentioned earlier, with the increased use of machine translation in industry settings and by professional translators, various industry tools such as PET246, TransCenter247 or iOmegaT (Aziz et al., 2012; Denkowski & Lavie, 2012; Moran et al., 2014) have been developed to measure productivity gains. However, only one tool, Inputlog, can, to my knowledge, register online activities. The online activities data can then be ‘infused’ into other logging systems such as CASMACAT or Translog II to combine text production data with the data on research activities (Carl et. al, 2016:41-42). Translog II has proven to be very useful in acquiring objective data about human translation-related processes such as reading and text production rhythms, mental memory, units of texts translators focus on etc. (Carl, 2012b); however, it is not a suitable tool for

246 http://rgcl.wlv.ac.uk/projects/PET/
247 http://www.findbestopensource.com/product/mjdenkowski-transcenter/
ecologically valid investigations of the translation process. First of all, Translog II is not a translation workbench, but a piece of software that records reading and writing processes. It therefore lacks all the functionalities translators are used to in their CAT tools and therefore cannot be used to study translation processes as they mostly happen in the workplace. Furthermore, the fact that it can only collect key strokes and eye movement limits the possible scope of investigation. Therefore, a tool designed to capture data directly from any environment would advance our knowledge about the translation process. An ideal tool to collect translation process data would benefit from having at least the following features:
- ‘pluggable’ to any existing translation or writing environment
- recording of keystrokes
- recording of eye movements
- recording of all interactions with online resources (automatically tied to the related research unit)
- recording of audio commentaries
- recording of translator’s facial expressions and body movements
- flexible note taking
- built-in interview and survey data tools
- built in translation QA tools
- coding
- export of data into qualitative and quantitative analysis software.

Although this wish-list might be challenging to implement, it could certainly provide an inspiration for tool developers as it could be easily adapted to investigate other processes such as writing or editing.

To summarise, the findings of the present study provide the initial stimulus for the development of an intelligent research assistant that would streamline research activities within the translation process in accordance with the individual translators’ natural resource and research behaviour. Furthermore, it provides a stimulus for future ecologically valid investigations into the online behaviour of professional translators and for further inquiry into resource and research styles and their attendant taxonomies (RTUT and TTRS). It is hoped that the present research has highlighted the need for the broadening of the scope of investigations into the translation process in terms of methodological underpinnings, conceptual frameworks and interdisciplinary efforts; it has called for a broader investigation of translator-computer and translator-information interactions in the ever changing socio-technological and techno-cognitive contexts. With this need come practical requirements for the development of tools and resources that would support further research into the translation process enabling process-oriented studies to be scaled-up and a sound scientific body of knowledge about the translation process to be produced. By doing so, perhaps we can come closer to being in a position to become “not only a borrower but also a lender” (O’Brien, 2013:13) to other disciplines.

8.8. Summary of contributions

The present study offers numerous original contributions to translation studies, especially to the better understanding of the translation process, with a particular focus on research-related activities carried out in an online environment. It addresses previously empirically unexplored or underexplored areas of the translation process which, with a few exceptions, have been mostly investigated as part of translation activities (i.e. interactions with the source and target text) rather than with a specific focus on the bursts of online research activity aimed at solving problems encountered in the ST and/or the TT. One of the main
assumptions of this thesis was that although research activities undoubtedly constitute an integral part of the translation process, they deserve to be studied in their own right. Despite the difficulties in separating the interactions with resources from interactions with the texts (ST and TT), the present study arguably succeeded in isolating online research activities to a sufficient degree and thus corroborated this assumption.

8.8.1. Systematising patterns of interaction between translators and online resources

The main contribution of the present study is the systematisation of the interaction patterns found between professional translators and online resources and the formulation of two taxonomies:

- Resource Type User Taxonomy (RTUT) representing the resource style
- Taxonomy of Translator Research Styles (TTRS) representing the research style.

The Resource Type User Taxonomy (RTUT) offers a new perspective on how translators use online resources. An innovative approach to conceptualising this use is suggested by considering the way(s) in which such resources are accessed (top-down vs bottom-up), the dominant orientation (ST vs. TT) as well as the general nature of the resources (termino-lexicographic vs. text-based). This provides novel insights into translators’ resource styles revealing individual preferences and some of the factors driving the translators’ choices.

The Taxonomy of Translator Research Style (TTRS) offers a new way of conceptualising overall translator research activities. It supports the identification of a variety of characteristic online behaviours and their exploration from a quantitative and a qualitative perspective, corresponding to the primary and secondary categories of TTRS respectively. The categorisation of the 16 participants into 5 types (Explorative, Prolific, Methodical, Understated and Economical) constitutes a new way of framing the complex and multidimensional nature of research activities into a simplified matrix, but at the same time communicates the high granularity of the analysed phenomena which cumulatively represent an individual research style.

8.8.2. Novel methodology to study patterns of research behaviour

This systematisation of the above mentioned patterns became possible through the development of a methodology to discern, analyse and study these patterns. Firstly, the cascading approach to data collection (the broad bird’s-eye view nature of the Global Survey followed by the much more specific, deep, explorative nature of the Main Study) can be seen as a promising model for process-based studies of this kind.

Secondly, the present study provides an analytical framework against which research activities can be analysed. This includes a framework for identifying resource styles (although it has been acknowledged that this framework hinges upon the current state of technology and further developments related to the convergence trends in technology) and the more comprehensive framework for the identifying research styles. The unique approach to the categorisation of the observed online behaviour also provides a notable contribution. This categorisation involved firstly, dividing the categories into primary and secondary, and secondly, the treatment of primary categories as relating to quantifiable time and volume aspects of translators’ research activities and the secondary categories as relating to qualitative aspects of these activities. In addition, the adaptive nature of this framework presents opportunities for its further development, refinement and improvement.
8.8.3. Contribution to workplace studies

The wealth of data gathered through this research provides rich and multidimensional insights into the usage of external resources during translation (and post-editing) processes, offering an invaluable descriptive account of how a sample of translators actually works in the increasingly complex and heavily technologised environment of today and how they approach research tasks in these environments.

The novelty of this study also lies in the profile of the studied subjects. To the best of my knowledge, this is the first time the translation process of professional freelance translators working in their natural environment has been studied in depth using a multi-method approach. Thus, this study arguably contributes to the research community’s knowledge about the professional freelance translators’ workplace, in particular the aspects that relate to the use of tools and online resources.

8.8.4. Contribution to studies on translator online behaviour

One of the most revealing aspects of this study is that it has uncovered the various mechanisms which translators develop and apply during their interactions with online resources, thus contributing to the much under-researched area of translator online behaviour. The identification of a variety of these mechanisms, and a diverse range of information needs and behaviours of professional translators constitutes a significant contribution to our knowledge about the interactions between translators and online resources. The present study thus emphasises the need to study individual differences in these behaviours, rather than just common tendencies, which seems crucial especially in times of increased interest in the personalization of technology (O’Brien, 2016). The present research is, to my knowledge, the first of its kind to address these individual differences within a group of translators, with regard to the use of online resources and with a view to identifying whether these individual differences can be systematised.

8.8.5. Contribution to studies on the use of external resources in the translation process

Another added value of the present study relates to furthering our knowledge in the area of so-called ‘tool box’-oriented translation research through the examination of the use of external resources by two groups of translation professionals (N=540 and N=16) and through offering a quantitative glimpse into this use. The Global Survey provided quantitative insights into the use of resources by a large sample of translation professionals. It was complemented by a much deeper exploration of the use of resources in the Main Study, which produced granular and multidimensional findings with regard to the use of resources by translators.

Such knowledge includes awareness of the specific resources that are currently used by professional translators, the types of resources that are popular and other characteristics of these resources (as described in section 4.3.3.) that appeal to translators. This is particularly important to monitor in order to enable tool developers to provide the desired combination of features and their future offerings. Such knowledge is also important for any future research as a reference point in time, especially in the context of the rapid developments in this area and the fast-changing practices, habits and preferences of translators.
In addition to the knowledge of what resources are being used, a systematic classification of currently available resources, together with a description of state-of-the-art resources included in this Thesis, provides a similarly useful reference point and a valuable inventory of the current state of technology. This information offers a suggestion as to how the area of online external support can be conceptualised, thus being potentially beneficial to students of translation, tutors teaching research and documentary skills, as well as future researchers venturing into this area of study.

Lastly, the measurement of time spent on research activities is also an important empirical contribution to the understanding of the translation process as it draws attention to the fact that the average time is higher than one might expect based on casual or anecdotal evidence and that it can vary dramatically between individual translators.

**8.8.6. Benefits stemming from the adoption of an interdisciplinary perspective**

Another aspect of this research that can be perceived as having an added value is the broader, interdisciplinary perspective of information behaviour that it adopts. Subsequently, the framework can be used to study the information behaviour of other related professions, including writers, editors and any other profession which involves carrying out research tasks. This broader perspective allows the findings to be further studied by means of comparative analyses with the said related professions.

Furthermore, this broader perspective extends to making provisional links between research styles and other areas of study such as personality. Although aspects of personality have been studied in relation to translation, they have not been explicitly linked to research behaviour. Although the present study does not claim to study personality, some aspects of translator research behaviour have been shown to relate to this area.

Finally, in a much broader sense, the present study can be seen as contributing to research in situated, embodied and extended cognition. The translators’ interactions with online resources and the identified resource and research styles can be seen as part of the translators’ cognitive framework which extends beyond the brain, right into their working environments (and, inevitably, further beyond). In this sense, the research presented in this thesis constitutes a stepping stone towards a broader research agenda which opens up new avenues for embedding translation process research more firmly in the field of cognitive science.

**8.9. Limitations**

As any time- and resource-constrained research study, this one is not without its limitations. These limitations are of a varied nature; however, they mostly relate to methodological aspects of the study.
8.9.1. Ecological validity not fully achieved

The main limitation of the present study can be related to the notion of ecological validity, which has not been fully achieved. Despite the fact that it was intended to make this study as authentic as possible, certain aspects of the design had to be controlled in order to support the analytical framework.

One such element was the source text, which had to be the same for all participants in order to enable a comparative analysis. The observation of translators translating authentic texts would have been methodologically challenging from the point of view of the analysis and comparability (16 different source texts, different text types, different complexity of the source texts etc.), and potentially also from the confidentially point of view. Translators would have to seek permission from the clients and, frequently, the clients are not directly accessible. Furthermore, source texts can be of a confidential nature which would preclude them from being used in a research study.

Another aspect of ecological validity that was not fully achieved was the nature of the sample. It could be argued that in a truly ecological setting, the sample is a naturally occurring one. In the case of this study, this criterion would have been impossible to achieve bearing in mind the difficulties in recruiting participants, as described earlier. The only way to recruit an entirely naturally occurring sample would be to carry out this study in a translation company. However, this did not line up with the objectives of this research (to study professional freelance translators). Furthermore, getting access to a company creates a host of other issues, as meticulously described by Ehrenberger-Dow (2014), which could also have bearings on ecological validity.

8.9.2. The materials used in the study

In addition to the text being supplied to the participants, it could be argued that its length (412 words) was not sufficient to draw conclusions from the observed patterns. A more viable length would be about 2000-2500 words, which would have allowed an observation of an average days’ worth of activities. However, the analysis of the screen recordings of 16 days’ worth of screen recordings would have been an impossible task to achieve for a single researcher in the given timeframe. The cumulative time of the 16 screen recordings delivered by the participants in the present study amounted to 16 hours 2 minutes and 54 seconds. The analysis of these screen recordings was challenging enough as it involved manual separation of the research activities from translation activities, a transcript of all activities undertaken during the task as well as detailed account of the resource use, including the individual resources and the research path followed by each of the 16 participants.

Another possible limitation related to the materials used in the study is the fact that only one text was used. In process-oriented research studies, two texts of varying difficulty are often used in order to observe differences and draw comparisons. However, the decision to use only one text in the present study was motivated by the findings drawn from the pilot study. The comparison of the use of resources in two texts of varying difficulty yielded predictable findings (e.g. more unknown words entailed more world-related knowledge such as definitions) which did not reveal much in terms of the differences between the participants, whilst adding complexity by introducing another variable.
8.9.3. The number of participants

Another limitation of the present study could be linked to the number of participants in the Main Study. As mentioned in section 3.3.4., the recruitment of the participants proved extremely difficult and although a sample of 20 was envisaged, only a sample of 16 was achieved. Nevertheless, this is still a relatively large sample given that the study was carried out solely by one researcher without any external help with data collection or processing. As a comparison, a study undertaken by Enríquez Raído (2011) which is methodologically the closest to the Main Study of the present research had a sample of 8 (six student translators and two professional translators of varying levels of expertise), but it was not preceded with a survey, and help was available during the data processing stage.

It is possible, however, that the number of participants had an impact on the way the TTRS grid was populated (e.g. only one specimen can be found in two of the categories of TTRS - Prolific and Explorer). A further, larger-scale study would be needed to establish whether the TTRS grid is an accurate representation of the spectrum of online behaviour observed on the screen relating to the observed timings and volumes or whether the categories/attributes related style need to be further investigated and validated.

8.9.4. The use of statistical methods

The lack of inferential statistics is a limitation that also has to be acknowledged as it prevents the results of this study from being generalised from the sample to the larger population of professional translators. This is certainly a limitation that could be addressed in further research. However, it was considered that the explorative and partially qualitative nature of the Main Study would not lend itself to the extensive use of inferential statistics, although a set of calculations (e.g. z-scores to identify outliers) was performed where no methods from descriptive statistics were available. 3.2.3.
REFERENCES


Heinström, J. (2006b). Fast surfing for availability or deep diving into quality - motivation and information
seeking among middle and high school students. Information Research, 11(4).


LinkedIn. (2014). Retrieved from https://www.linkedin.com/groupItem?view=&gid=4168755&type=member&item=5934964395703570434&commentID=5935772881530138624&trk=groups_guest_most_popular-0-b-cmn&goback=gmp_4168755


Roget, P. M. (1853). Thesaurus of English words and phrases, classified so as to facilitate the expression of ideas (2nd ed.). London: Longman, Brown and Green.


279


TAUS. (2011). The future for translators looks bright, but they will have to reinvent the profession first.
TAUS. Retrieved from https://www.taus.net/think-tank/articles/event-articles/the-future-for-translators-looks-bright-but-they-will-have-to-reinvent-the-profession-first


LIST OF APPENDICES

Appendix 1
Main Study Post-task e-mail questionnaires

Appendix 2
Source texts used in Self-study, Pilot Study and Main Study

Appendix 3a
Global Study questionnaire

Appendix 3b
Main Study profile questionnaire

Appendix 4
Instructions for the Main Study

Appendix 5
Resources used in the Main Study translation task and flowcharts of research episodes

Appendix 6
The number of types of resources and the number of specific resources within the type accessed by individual translators in the Main Study

Appendix 7
Number of instances of specific resources use by individual translators showing the name of the resource

Appendix 8
Breakdown of the most used specific resources common to at least two participants

Appendix 9
Frequency of use of other types of resources by individual translators in the Main Study

Appendix 10
List of research units (RUs)