Consultation Behaviour with Online Resources in English-Chinese Translation: An Eye-tracking, Screen-recording and Retrospective Study

Yixiao Cui and Binghan Zheng[1]

School of Modern Languages and Cultures, Durham University, Durham, UK

ABSTRACT

This paper investigates the interaction between translators' perceived translation problems and their online consultation behaviours, and how different consultation behaviours affect translation acceptability. Previous studies indicate that online consultation includes various types of complex information-searching behaviours which, to a great extent, depend on the personal preferences of the web users. In this study, 38 MA translation students translated two 100-word texts from English (L2) into Chinese (L1) using Translog II, with their translations and consultation processes being registered by a Tobii TX300 eye-tracker. The main findings are as follows: (1) an increase in perceived translation difficulty leads to an increase in both the time spent on online consultation and the complexity of the consultation, but does not lead to an increase in the cognitive load expended on consultation; (2) general translation problems, which involve more resource types and longer search times, require more consultation time than specific translation problems; (3) two sub-types of consultation behaviour, information-seeking and information relevance evaluation, are purpose-driven and (4) longer consultation time results in higher acceptability of individual translation solutions, while higher consultation complexity does not.

KEYWORDS

Information behaviour, online consultation, eye-tracking, retrospection, English-Chinese translation

1. Introduction

With the development of information technology, online consultation resources have become increasingly important for professional translators at the workplace. Technologies that support translators in their work are evolving rapidly, with an increase in both quantity and quality (Fulford & Zafra, 2004; O'Hagan, 2012). Despite the fact that online resource consultation has been proved to make up as much as 25% of the translation process (Hvelplund, 2017), research on consultation during the translation process and how consultation affects translation quality remains at an early stage.

With focus group interviews of 19 professional translators, White, Matteson and Abels (2008) took an early step to explore the nature of consultation in translation. They provided a task-related information behaviour model in translation and called for further research on consultation in translation. The existing studies on the consultation process have had two different points of departure: (1) what is the nature of consultation and (2) how to improve

[1] CONTACT: Binghan Zheng, binghan.zheng@durham.ac.uk, School of Modern Languages and Cultures, Durham University, Elvet Riverside, New Elvet, Durham, DH1 3JT, UK

consultation accuracy. Using screen-recording, think-aloud protocols (TAPs), and post-task questionnaires, Gough (2016) summarized the patterns of professional translators' online resource consultation behaviours with two taxonomies: the Resource Type User Taxonomy and the Taxonomy of Translator Research Styles. Hvelplund (2017) studied 18 professional translators' behaviours in using digital resources during translation, and concluded that a considerable amount of attention (19.4%) was allocated to consultation, and consultation exerted a heavier cognitive load than translation drafting and revision. Based on the same dataset, Hvelplund (2019) further investigated the processing flow patterns involved in the consultation, and summarised four types of processing flow patterns, including ST-DR-ST (ST to digital resources to ST), ST-DR-TT, TT-DR-ST, and TT-DR-TT, with ST-DR-TT being the most common pattern. Trainee translators' consultation behaviours were also studied in some recent studies. Sycz-Opoń (2019) investigated the information-seeking behaviour of 104 translation students in legal translation using observation and TAPs. She concluded that translation students relied heavily on bilingual dictionaries, and behaved in conservative ways in source preferences and over-dependence on external resources. Shih (2019) studied 18 translation students' web searching behaviour using screen-recording and found that more extensive searches led to better search results. To answer the second question, several studies were conducted by comparing the behaviours of professional and trainee translators. Using direct observation via screen-recording, questionnaires and semi-structured interviews, Enríquez Raído (2011, 2014) studied the consultation behaviour of 6 participants with different levels of translation experience. She found that experienced translators used a wider variety of online resources and consulted background information more frequently than trainee translators.

Previous studies on consultation suggested that, as a complex process, consultation could be affected by various factors. However, few studies have provided a clear view of what factors might affect consultation and in what ways. In Information Science, however, web searching behaviours have been extensively investigated; those findings are germane to translation consultation research. Navarro-Prieto, Scaife and Rogers (1999) studied web searching tasks under different situations and developed an Interactive Framework to explain the interaction between task and information searching behaviour. Based on the results, they identified three types of cognitive strategies used in web searching: (1) the top-down strategy, which is when users search in a general area and then narrow down their search from the links provided until they find what they are looking for; (2) the bottom-up strategy, which is when users look for a specific keyword, and (3) the mixed strategy, which is a combination of the above two strategies. This categorisation of search strategies was used by Gough (2016) in her study of professional translators' consultation behaviours. Using questionnaire data White and Iivonen (2001) studied the influence of question-related variables on web users' choice of an initial search strategy and developed a model to explain the selection process. These studies confirmed that users' web searching processes varied according to the type of information they required. It, therefore, seems reasonable to assume that consultation behaviour in translation varies according to the type of translation problem that has been encountered and that an investigation of the correlation between the two aspects would be worthwhile. Another important aspect of research into web searching is the investigation of how users evaluate information. Granka, Joachims and Gay (2004) used eye-tracking data to study the behaviour of users when reading abstracts and evaluating links before making the final click. They compared the amount of attention paid to each link and found that users tended to pay much more attention to the first and second links than to the rest.

Although previous studies have investigated online consultation in translation to some extent, the complex relationships among sub-types of consultation as information-searching behaviour remained unclear. Therefore, we posed the following questions: (1) What is the

effect of perceived translation difficulty on consultation behaviour? (2) How does the type of translation problem relate to information-searching behaviour and the evaluation of information relevance? (3) How does consultation behaviour affect the acceptability of translation solutions?

2. Information behaviour

The term 'Information Behaviour' is commonly used in the discipline of Information Science. It refers to 'activities a person may engage in when identifying his or her own needs for information, searching for such information in a way, and using or transferring that information' (Wilson, 1999, p.249). In translation, consultation behaviour is triggered by translation problems, involves searching for information, and is used as an aid in producing the target text; it can therefore be viewed as a kind of information behaviour.

Several models have been proposed to explain information behaviour. Ellis (1989) developed a model based on studies of the information seeking characteristics of social scientists; these could be simply described in words in the form of a set of theoretical propositions. Derived from a series of five studies investigating users in information-seeking situations, Kuhlthau (1991) developed a six-step model with users as the focus. Based on previous models, Wilson (1999) proposed a linear model (Figure 1) to define clearly each step in information behaviour, covering both users and the search system, and this was employed in the present research.

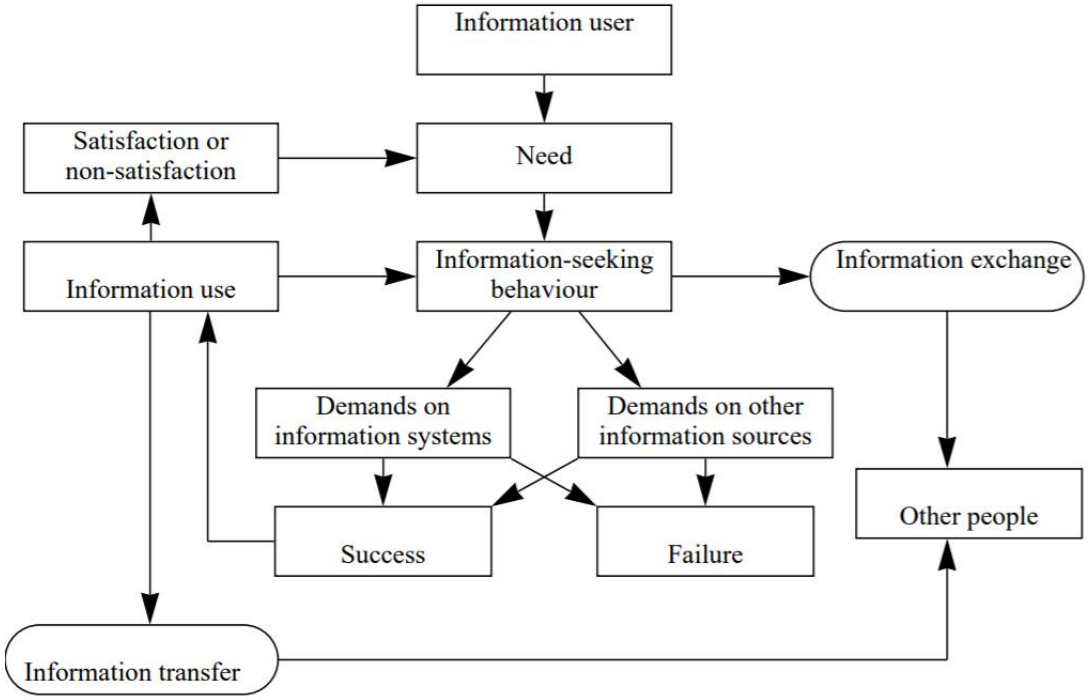

Figure 1. Wilson's model of information behaviour (Wilson 1999, p.252).

According to Wilson (1999), most information behaviour models have similar features: 'they are statements, often in the form of diagrams, that attempt to describe an information-seeking activity, the causes and consequences of that activity, or the relationships among stages in information-seeking behaviour' (p.250). For our investigation of consultation in

translation, we simplified Wilson's model into three steps: information need, information-seeking behaviour, and information use behaviour (see Figure 2).

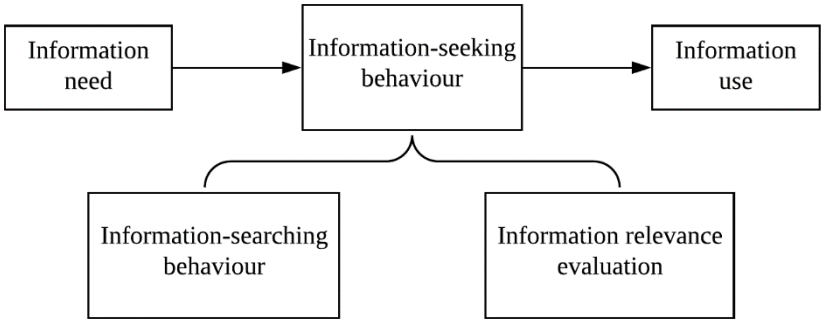

Figure 2. Simplified Wilson's model of information behaviour.

2.1 Information need

Miranda and Tarapanoff (2008) defined information need as 'a state or process started when one perceives that there is a gap between the information and knowledge available to solve a problem and the actual solution of the problem' (p.2). For all kinds of information behaviour, information need marks the start of the entire process, since subsequent information-seeking behaviour is 'prompted by the individual's physiological, cognitive or affective needs' (Wilson, 2000, p.52). Angelone (2010) described translation as 'a chain of decision-making activities relying on multiple, interconnected sequences of problem-solving behaviour for successful task completion' (p.17). As a problem-solving process, the first step in translating a text is problem identification. This triggers the subsequent procedures of information behaviour,[2] which might include the consultation of online resources.

In order to investigate the interaction in translation between information need and information-seeking behaviour, we applied Angelone's (2010) categorisation which includes: (1) source language comprehension, which refers to the inability to understand the source text; (2) source language-target language transfer of meaning, which refers to difficulty in finding a satisfactory equivalent in the target language, and (3) target language text production, which is connected with the style of, or cultural references in, the source text. For better comprehension, these three types of translation problems are referred as Type 1, Type 2, and Type 3 respectively in this article. We employed Angelone's categorisation in this research, since it covers all types of translation problems from the translators' behaviour aspect, and clearly describes the different types of information need at each step.

One of the major features of information need is uncertainty (Wilson, 2000). Different translators may have different types of problems with the same text. Therefore, we applied retrospective interviews soon after the translation tasks to support data analysis relating to problem categorisation.

2.2 Information-seeking behaviour

Information-seeking behaviour is 'the purposive seeking for information as a consequence of a need to satisfy some goal' (Wilson, 2000, p.49). Due to the nature of this study it refers to consultation with online resources only, with no access to off-line resources allowed. Online

[2] This term was also used by Enríquez Raído (2011, 2014), but with a different definition which is equal to "translation problem" as defined in this study.

resources are defined as external internet-based aids that are not integrated into the text processing interface, and are differentiated from CAT tools such as SDL Trados and MemoQ (Hvelplund, 2017, p.72).

Information-seeking behaviour is a complex process that comprises various sub-types, including information-searching behaviour and information relevance evaluation (Wilson, 2000, p.49). Screen-recording and eye-tracking data were used in our study, with the former indicating the number of queries and resource type, and the latter indicating the amount of cognitive effort involved in individual segments. In order to facilitate the further analysis of information relevance evaluation, 'online information' was divided into three categories: (1) lexical information, which contains information related to the meaning of words (i.e. definitions and translations provided in dictionaries); (2) extra-lexical information, which contains all other linguistic-related information (i.e. grammatical information and synonyms), and (3) extra-linguistic information, which contains non-linguistic related information (i.e. background information). Salojärvi, Kojo, Simola, and Kaski's (2003) usability test indicated that information relevance could be measured by total fixation duration, which is therefore used as a measurement to indicate perceived information relevance in our study.

2.3 Information use behaviour

Information use behaviour has been defined as 'the physical and mental acts involved in incorporating the information found into the person's existing knowledge base' (Wilson, 2000, p.50). In this paper, it refers to the production of a target text with the support of both internal knowledge and the results of external consultation. To investigate information use, a product-oriented approach was adopted in our study, with translation acceptability being used to assess the effectiveness of consultation. Since the acceptability assessment of an entire target text might be affected by elements other than consultation, such as translation competence or personal working habits, in order to obtain a more accurate evaluation of the effect of consultation on translation acceptability, only the acceptability of individual translation solutions for which consultation had been used was assessed in this study.

3. Research Design

3.1 Participants

38 MA translation students (32 females and 6 males) with an average age of 24 years (range 21-35 years, *SD*=2.67 years) were recruited as participants on a voluntary basis. They were all native Mandarin Chinese speakers with English as their second language, and had learned English from an average age of 8.63 years (range 3-13 years, *SD*=2.33 years). The participants had a mean IELTS score of 7.38 (range 7-8, *SD*=0.36). None of them had been brought up in a bilingual environment or had worked as a professional translator before. They were all touch-typists and had normal or corrected-to-normal vision. To minimise any negative influences on data quality, the participants were asked not to drink coffee or alcoholic beverages before the experiment, and not to wear heavy mascara during the experiment. They were told that anonymity and confidentiality would be ensured, signed a consent form and were rewarded with a supermarket gift card. The experiment was approved by the research ethics committee of Durham University.

3.2 Materials

In line with Enríquez Raído (2014), we used two source texts with different perceived difficulties: Text A and Text B (see Appendix I). Text A was an excerpt from an article published in *New Scientist*, a weekly English-language magazine that covers all aspects of science and technology. Text B was an excerpt from *Coral Reef and Global Climate Change*, a popular science book introducing the negative effects of global climate change on coral reefs. The texts were of similar length in terms of the total number of words. Following Jensen (2009), three factors - readability,[3] percentage of complex words and non-literalness - were used to measure the text complexity (see Figure 3). In average, the results indicate that Text A requires 8.68 years of schooling for successful comprehension, and Text B 18.02 years. The percentage of complex words indicates that Text B contains a larger proportion of complex words than Text A. Combining the results of all the three indicators, Text B is assessed as more complex than Text A.

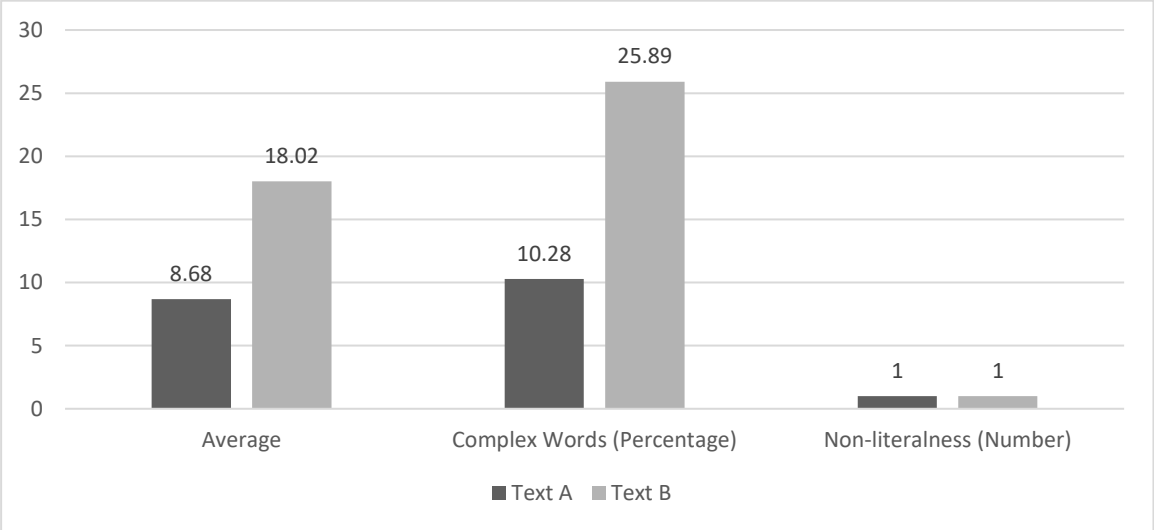

Figure 3. Summary of text complexity by three indicators.

We recruited 12 freelance translators to rate the translation difficulty of the two texts on a five-point Likert scale, with 1 as 'very easy' and 5 as 'very difficult'. The mean value of translation difficulty for Text A was 2.18 (range 1-3, *SD*=0.72), and for Text B 4.27 (range 3-5, *SD*=0.58). The result of a paired t-test showed a statistically significant difference between the two sets of numbers ($t(11)= -9.381$, $p<.001$, $d=3.07$). In brief, Text B was more complex and perceived as more difficult to translate than Text A.

3.3 Experiment settings

The experiment was prepared and run in an eye-tracking lab. All the participants' eye movement data were recorded with a desktop version of the Tobii TX300 eye-tracker. The eye-tracker was connected to a 23" LCD monitor that functioned as the presentation screen. The screen resolution was set at 1280*1024 pixels and the fixation radius was the default setting of the Tobii system, 35 pixels per inch.

To suit the eye-tracker based design, we purposely split the screen into two equal areas (shown in Figure 4), with the Translog II user interface on the left for translation and the web

[3] The readability indices include the Automated Readability Index (ARI), the Flesch-Kincaid index, the Coleman-Liau index, the Gunning Fog index, the SMOG index, the Flesch Reading Ease Score index, and LIX. Results of the readability indices and the percentage of complex words can be retrieved from https://www.webfx.com/tools/read-able/.

browser (Internet Explorer 11) on the right for consultation. The English source texts were displayed in the upper window of the Translog interface,[4] with a Microsoft Sans Serif typeface set at 18 points, and triple line spacing.[5] The Chinese target texts were produced in the lower window, with a SimSun typeface set at 18 points, also with triple line spacing. The web browser was set on a blank page before the translation task began. After each task, the search history was erased in order to avoid any potential influence on the next participant.

Figure 4. Interface design for the experiment.

3.4 Experiment procedure

Being tested individually, the participants were firstly asked to sit approximately 60cm away from the monitor; then a five-point calibration and validation procedure was carried out. After an acceptable calibration had been saved, each participant started to translate the warm-up text and then the two experimental texts with no time constraints. The participants were allowed to use any online resources apart from machine translation and CAT tools. They could take a break between tasks if necessary. After translating the texts, they were asked to take part in a retrospective interview to enable us to categorise their translation problems and obtain more information about their consultation behaviour. Finally, they were asked to complete a questionnaire about their educational backgrounds and their familiarity with background knowledge for the two source texts (see Appendix II). The total session for each participant lasted roughly 90 minutes.

4. Results

4.1 Quality assessment of eye-tracking data

[4] The key-logging data from Translog II were not analysed in this paper. Translog II was used as the interface to display the source text and produce the target text.

[5] We set up triple line spacing in order to register more accurate eye-tracking data with text processing, and to minimise the possible deviation caused by the restriction of spatial resolution.

The quality of the eye-tracking data collected was assessed prior to the data analysis. First, the overall quality of the eye-tracking data obtained for each translation task was assessed. With reference to Hvelplund (2014), the following three criteria were adopted in this study for assessment: Mean Fixation Duration (MFD), Gaze Time on Screen (GTS) and Gaze sample to Fixation Percentage (GFP). MFD is calculated as total fixation duration divided by the number of fixations. In this study, MFDs below 206.85 ms (one SD below the mean) were considered as invalid data. GTS is a calculation of total fixation duration as a percentage of total task time [(total fixation duration/total task time)*100%]. GTS data below 46.60% (one SD below the mean) were considered as invalid data. GFP is calculated by comparing the total number of gaze samples with the total number of gaze samples that formed part of a fixation [(number of gaze samples/number of fixation gaze samples)*100%]. GFPs below 74.67% (one SD below the mean) were considered as invalid data. Data that satisfied at least two out of the three criteria were included for further analysis. As can be seen from Table 1, the data obtained from P_5, P_{10}, P_{12}, P_{16}, P_{21}, P_{25}, P_{30} and P_{38} were deemed invalid, with the percentage of invalid data being 21.05%.

Table 1. Summary of eye-tracking quality assessment with invalid data (marked as ×).

Participant	**Text A**			**Text B**		
(P_n)	**MFD**	**GTS**	**GFP**	**MFD**	**GTS**	**GFP**
P_5	×	×	×	×	×	×
P_{10}	×		×	×		×
P_{12}		×	×			×
P_{16}	×	×			×	
P_{21}	×	×	×	×	×	×
P_{25}	×	×	×	×	×	×
P_{30}	×			×	×	
P_{38}				×	×	

The next step was to assess the quality of eye-tracking data for each type of individual consultation behaviour. We manually selected every individual consultation behaviour from the entire translation process, starting from typing the keyword in the internet browser and continuing until the start of the next search. In some cases, if a participant conducted noncontinuous searches about the same source text segment, it would be counted as one consultation behaviour.

For the eye-tracking quality of individual consultation behaviours, two conditions needed to be met in the data analysis: (1) at least one fixation had been recorded on the consultation section to be used for calculating cognitive effort; (2) the three different information types (lexical, extra-lexical and extra-linguistic) could be clearly distinguished from each other so that the correlation between the consultation and the evaluation of information relevance could be analysed. In this research, 481 individual consultation behaviours met the two standards and were considered as valid data for further analysis.

4.2 Effect of translation difficulty on consultation behaviour

The interaction between translation difficulty and consultation was studied from three aspects: (1) the amount of visual attention distributed on consultation, indicated by the percentage of total fixation duration (seconds) and fixation count on consultation over the entire task; (2) cognitive effort, indicated by mean fixation duration; and (3) consultation complexity, indicated by the number of resource types. We hypothesised that an increase in translation

difficulty perceived by participants would lead to an increase in more visual attention, heavier cognitive effort distributed on consultation, and more types of online consultation resources.

In our experiment design, there were no time constraints, which means that the length of time spent on consultation could vary depending on the participants' personal processing behaviours, and the raw data would not reveal the actual correlation between translation difficulty and the amount of cognitive effort devoted to different processing phases. Therefore, instead of using the raw data of total fixation duration and fixation count, we used the percentage of total fixation duration and fixation count allocated to consultation over the entire task. Table 2 shows the mean value of the proportion by 30 participants. The differences in the percentage of total fixation duration and fixation count between the translations of the two texts were both statistically significant (paired t-test, $t(29)= -6.844$, $p<.001$, $d=1.31$; $t(29)=-6.989$, $p<.001$, $d=1.42$).

Table 2. Mean value of the percentage of total fixation duration and fixation count on consultation.

Text	N	Total Fixation Duration (percentage)		Fixation Count (percentage)	
		Mean	SD	Mean	SD
Text A	30	15.12%	8.92%	13.85%	8.38%
Text B	30	26.92%	9.15%	26.18%	9.01%

Table 3 shows the mean value of fixation duration (milliseconds) on processing the source text, the target text and online consultation resources. The results show that mean fixation duration on consultation is almost the same for the translation of Text A and B. The difference in the mean fixation duration on consultation between the translations of the two texts was not statistically significant (paired t-test, $t(29)=0.309$, $p>.05$).

Table 3. Mean Fixation Duration on processing ST, TT and consultation.

Text	N	Mean Fixation Duration (ms)		
		ST	TT	Consultation
Text A	30	217.03	266.83	265.84
Text B	30	234.98	276.67	263.71
Total	60	226.00	270.25	264.68

PACTE's categorisation of online resources was used to calculate the number of resource types (Kuznik & Olalla-Soler, 2018, pp.31-32). The list contains 8 items: (1) search engines, such as *Google* or *Baidu*; (2) bilingual dictionaries, such as *Youdao Dictionary* or *Bing Dictionary*; (3) monolingual dictionaries, such as *Oxford English Dictionary*; (4) dictionaries of synonyms, such as *WordReference*; (5) encyclopaedias, such as *Wikipedia*; (6) databases, such as *UNTERM*; (7) online corpora, such as *Collins*; and (8) online or field-specific portals, such as information related to the subject of the source texts. In this research, only six types (excluding dictionaries of synonyms and databases) of online resource were used by the participants. Table 4 shows the mean number of resource types in translating Text A and Text B, with the difference between the two sets of figures being statistically significant (paired t-test, $t(29)= -3.710$, $p<.05$, $d=0.83$).

Table 4. Mean number of resource types in translating Text A and Text B.

Text	N	Number of Resource Types	
		Mean	SD
Text A	30	1.40	0.89
Text B	30	2.27	1.17

4.3 Correlation between translation problems and information-seeking behaviour

As mentioned in section 2.1, translation problems are categorised into three types: comprehension problems, transfer problems and production problems. Overall, 481 translation problems were categorised (see Table 5).

Table 5. Number and percentage of three types of translation problems.

Text	Type 1		Type 2		Type 3	
	N	Percentage	N	Percentage	N	Percentage
Text A	86	59.72%	31	21.53%	27	18.75%
Text B	181	53.71%	96	28.49%	60	17.80%
Total	267	55.51%	127	26.40%	87	18.09%

As mentioned in section 2.2, information-seeking behaviour can be further divided into information-searching behaviour and information relevance evaluation. The results for the correlation between consultation and the two sub-types are presented below.

4.3.1 Information-searching behaviour

The correlation between translation problem type and information-searching behaviour was investigated using four measurements: (1) total fixation duration(s); (2) fixation count; (3) number of queries, defined as the number of webpages visited in one consultation; and (4) number of resource types. The first two measurements were used to indicate the amount of cognitive effort; while the last two were used to indicate the consultation complexity. Due to the nature of each type of translation problem (see section 2.1), we proposed that the information-searching behaviour for consulting Type 1 problems would require the least cognitive load, hence have the lowest consultation numbers and resource types.

Table 6 shows the mean value of the four measurements for the three different types of translation problem. All four measurements reveal an upward trend from Type 1 to Type 3 problems.

Table 6. Mean values of the four measurements for the three types of translation problem.

Translation Problem Type	N	Total Fixation Duration(s)		Fixation Count		Number of Queries		Number of Resource Types	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Type 1	267	4.43	4.16	18.36	17.05	1.10	0.38	1.06	0.27
Type 2	127	11.17	14.36	45.30	58.08	1.54	1.10	1.24	0.56
Type 3	87	18.56	24.27	80.39	98.21	1.91	1.44	1.36	0.70

Since the homogeneity of variance of this dataset was violated, nonparametric tests (Kruskal -Wallis tests) were conducted to show the pairwise comparisons (see Table 7). Apart from the comparison between resource type number used for Type 2 and Type 3 problems, all the other comparisons demonstrated statistical significance.

Table 7. Results of between-group comparison for the four indicators.

Comparison Pair	Measurement	*p*-value
Type 1 vs. Type 2	Total Fixation Duration	<.001
	Fixation Count	<.001
	Number of Queries	<.001
	Number of Resource Types	<.001

Type 1 vs. Type 3	Total Fixation Duration	<.001
	Fixation Count	<.001
	Number of Queries	<.001
	Number of Resource Types	<.001
Type 2 vs. Type 3	Total Fixation Duration	<.001
	Fixation Count	<.001
	Number of Queries	<.05
	Number of Resource Types	>.05

To further assess the correlation between translation problem type and consultation, a Spearman correlation coefficient was conducted and showed a significant correlation for all the four indicators (see Table 8).

Table 8. Spearman correlation coefficients between translation problem type and consultation.

	Translation Problem Type	
	Rho	***p*-value**
Total Fixation Duration	0.440	<.001
Fixation Count	0.448	<.001
Number of Queries	0.366	<.001
Number of Resource Types	0.259	<.001

4.3.2 Information relevance evaluation

The correlation between translation problem type and information relevance evaluation behaviour was investigated by calculating the proportion of attention distributed on each type of information, including lexical information, extra-lexical information, and extra-linguistic information (see section 2.2). By drawing different Areas of Interest (AOIs) in the consultation interface with Tobii Studio (the eye-tracking software for Tobii TX300), we were able to calculate the total fixation duration allocated to processing each type of online information in each individual consultation behaviour. A higher total fixation duration indicated that the participant spent more cognitive effort on one information type than on the others and perceived it to be more relevant for solving the translation problem. We hypothesised that from Type 1 to Type 3 translation problems, the proportion of attention allocated to lexical information would decrease while that allocated to the other two types of information would increase.

The mean percentages of attention allocated to each type of information for the three types of translation problems are presented in Table 9.

Table 9. Mean values for attention allocated to each type of information.

Translation Problem Type	**Lexical Information**	**Extra-lexical Information**	**Extra-linguistic Information**
Type 1	79.57%	16.46%	3.97%
Type 2	57.15%	27.24%	15.61%
Type 3	39.13%	34.56%	26.31%

A Spearman correlation coefficient was conducted to further assess the correlations between translation problem type and the proportion of cognitive effort allocated to the three types of information. The results showed significant correlations for all three information types (see Table 10).

Table 10. Spearman correlation coefficients between translation problem type and cognition allocation to three types of information.

	Translation Problem Type	
	Rho	***p*-value**
Lexical Information	-0.421	<.001
Extra-lexical Information	0.240	<.001
Extra-linguistic Information	0.318	<.001

4.4 Correlation between consultation and translation acceptability

As indicated in 2.3, we only assessed the acceptability of translation solutions for which online resources had been consulted. PACTE's (2009) acceptability assessment criteria were applied in this study; these are grouped into three main categories: meaning (i.e. the extent to which source-text meaning is reproduced), function (i.e. how adequately the function of the translation and the translation brief have been achieved) and language (i.e. how appropriate the use of the target language is). An assessment result was given for each category, and the operational definitions are presented in Table 11.

Table 11. Operational definitions for the three assessment results (PACTE, 2009).

Result	Value	Operational Definition
Acceptable solution (A)	1	The solution activates all the relevant connotations of the ST in the TT.
Semi-**A**cceptable solution(SA)	0.5	The solution activates some of the relevant connotations of the ST in the TT, and maintains the coherence of the TT.
Non-**A**cceptable solution (NA)	0	The solution activates none of the relevant connotations of the ST in the TT, or introduces connotations that are incoherent

27 possible permutations existed for each solution, which were finally considered acceptable (A), semi-acceptable (SA) or non-acceptable (NA) according to the scheme presented in Appendix III, with each category being given a numeric value. Each translation solution was reviewed and assessed by two reviewers who were lecturers in Translation Studies at a UK university. Where doubts arose, both reviewers were consulted to agree on a result. In total, 481 translation solutions were assessed, and the results are shown in Table 12.

Table 12. Number and percentage of assessments on individual translation solutions.

Translation Problem Type	N	Assessment					
		A		SA		NA	
		N	Percentage	N	Percentage	N	Percentage
Type 1	267	200	74.91%	34	12.73%	33	12.36%
Type 2	127	84	66.14%	25	19.69%	18	14.17%
Type 3	87	69	79.31%	7	8.05%	11	12.64%
Total	481	353	73.39%	66	13.72%	62	12.89%

A Spearman correlation coefficient was conducted to investigate the correlation between consultation and the acceptability of each individual translation solution using three measurements: (1) total fixation duration (s); (2) fixation count; (3) number of queries, and (4) number of resource types. We proposed that the increase in consultation length and complexity would lead to an increase in translation acceptability.

The correlations between the first two measurements (total fixation duration and fixation count) and translation acceptability were found to be significant, but no significant

correlations were found between the other two indicators and translation acceptability (see Table 13).

Table 13. Spearman correlation coefficients between consultation and translation acceptability.

	Translation Acceptability Assessment	
	Rho	***p*-value**
Total Fixation Duration(s)	-0.122	<.05
Fixation Count	-0.105	<.05
Number of Queries	0.056	>.05
Number of Resource Types	0.019	>.05

5. Discussion

5.1 Effect of translation difficulty on consultation

Vakkari (1999) provided a model (see Figure 5) to explain the relationship between task complexity and information behaviours. According to this model, task complexity is determined by "the degree of predeterminability of task performance" (p.826), with one of the central features of predeterminability being "its information requirements" (p.826).

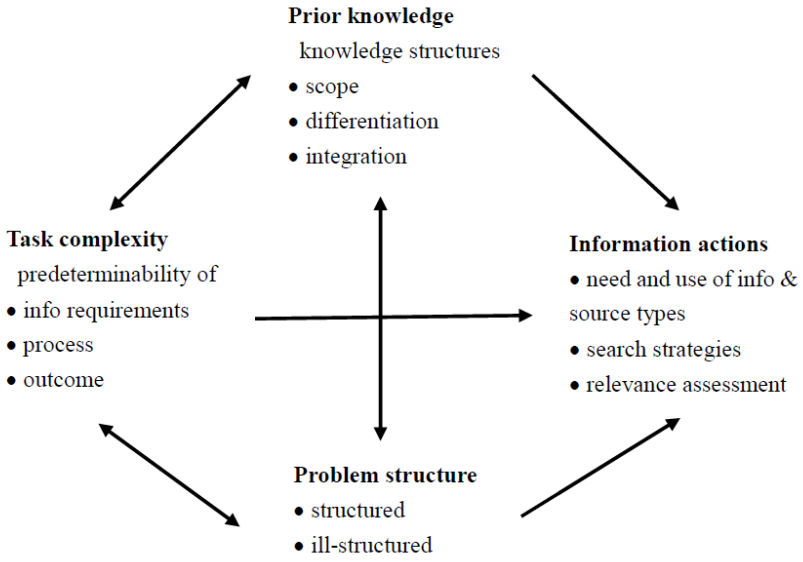

Figure 5. Task complexity and information behaviours model (Vakkari, 1999, p.830)

Based on Vakkari's model, since Text B was perceived as more difficult to translate than Text A, we could hypothesise that translating Text B would require more consultation information than translating Text A. According to Wilson's model (Wilson, 1999, p.252), information behaviour is triggered by the gap between a subject's prior knowledge and the knowledge required to solve the task. As the gaps for the two translation tasks were different, we assumed that consultation would also be different in terms of the translators' behaviours and the cognitive resources required to fill the gaps. The effect of translation difficulty on consultation behaviour was investigated from three aspects: attention distribution, consultation complexity and cognitive effort.

The percentage of total fixation duration and the fixation count allocated to consultation was used to indicate the proportion of consultation over the entire translation process. In some

previous studies, this proportion was calculated using the length of consultation in time (Daems et al, 2016; Hvelplund, 2017; Kuznik & Olalla-Soler, 2018). However, our observation of the translation process revealed that the participants' attentions flitted frequently between translation and resource consultation, which made it difficult to assess precisely the length of the consultation. Hvelplund (2017, 2019) used the percentage of gaze time to investigate the attention distribution on consultation in the translation process. Similarly, in this research, we used the percentage of total fixation duration and the fixation count allocated to consultation over the entire translation process to calculate the proportion of consultation, and found that both indicators were significantly higher in translating Text B than Text A. The complexity of consultation was indicated by the number of resource types consulted. When translating Text B, participants used more types of online resources that that in translating Text A. A similar result was reported by Enríquez Raído (2014), that the increase of perceived task complexity led to a wider range of search behaviour, including using more types of resources. These results indicate that for more difficult translation tasks, both the amount of cognitive effort devoted to consultation and the complexity of consultation behaviour show an upward trend.

However, more attention distributed to consultation did not mean that consultation for a more difficult task was more cognitively consuming. Based on a review study, Rayner (1998) concluded that mean fixation duration was different for different tasks, and pointed out that it could be used as an indicator for cognitive effort. Liu, Zheng, and Zhou (2019) also found mean fixation duration increased in translating more difficult texts. In this research, we compared mean fixation duration on consultation in translating Text A and Text B and found no significant difference. In other words, although the translation difficulty increased, participants did not expend greater cognitive effort on consultation. The entire consultation process consisted of individual information-searching behaviours, so the cognitive load devoted to consultation should be determined by the cognition expended on individual behaviours. In this study, the proportions of each type of translation problems were similar to each other (see Table 5 in Section 4.3), so the proportions of each type of information-seeking behaviour should also be similar. As a result, although the higher translation difficulty creates a greater information need in translating Text B, the increase is more in consultation quantity, not in consultation complexity. Therefore, the increase in translation difficulty does not affect the cognitive effort expended on consultation behaviour.

5.2 Correlation between translation problems and information-seeking behaviour

In this research, we categorised all the translation problems into three types based on the translation process. Since this categorisation was only related to individual translation problems, the effect of perceived translation difficulties was not considered. Based on the nature of each type of translation problem (see section 2.1), we could further categorise the information-seeking behaviours for three types of translation problems as specific fact-finding tasks, extended fact-finding tasks and exploratory tasks, respectively (see Table 14). This categorisation was presented by Shneiderman (1997).

Table 14. Categorisation of three types of translation problems.

	Definition	**Information Task Type**
Type 1	Comprehension of the source language	Specific fact-finding task
Type 2	Production of satisfactory equivalence	Extended fact-finding task
Type 3	Production of target language	Exploratory task

Navarro-Prieto, Scaife, and Rogers (1999) proposed an Interactive Framework, which can be used to explain web users' searching strategy. According to this framework, web users'

searching strategy is purpose-driven. For fact-finding tasks, users search with a specific keyword (bottom-up strategy), while for exploratory tasks they first search in a general area and then narrow down the search results (top-down strategy). A similar correlation was found between translation problem type and consultation behaviour in this research. As the information need became less specific, both the length and the complexity of the participants' information-searching behaviour increased.

This correlation could also be indicated by the attention allocated to different types of information. For Type 1 problems, the information-seeking behaviour tended to be focused on answering a specific question; this type of problem produced the lowest number of queries among the three types, and the number of resource types consulted was also lower than for the other two types of problem. Similarly, when evaluating information relevance for Type 1 problems, the participants devoted a large proportion of their attention to one type of information – lexical information (accounting for 79.57% on average). For Type 2 and Type 3 problems, information-seeking behaviour became more general, with no specific target being set, which can be indicated by the increased number of queries and the increased number of resource types used.

5.3 Correlation between consultation and translation acceptability

Since a large proportion of working time is spent on consultation during translation, it seems reasonable to assume that the use of online resources should bring about an increase in translation acceptability. However, previous studies have refuted this hypothesis by indicating that neither the length nor the complexity of consultation is correlated with translation acceptability (Daems et al, 2016; Kuznik & Olalla-Soler, 2018). Kuznik and Olalla-Soler (2018) explained their results by proposing that translation acceptability was determined mainly by students' internal support, which was their 'linguistic and extra-linguistic knowledge, their knowledge of translation and cognitive strategies' (p.24). External resources served only as supplements but were not decisive elements.

However, since we have highlighted the effect of consultation on translation acceptability by only evaluating individual segments, we have found different results. A significant correlation was found between consultation length and the acceptability of individual translation solutions. A longer consultation led to a higher level of translation acceptability. This finding indicates that online resources are effective aids for trainee translators to increase their translation acceptability. On the other hand, no correlation was found between the complexity of consultation and translation acceptability. Similar results were reported in previous studies. Olalla-Soler (2018) found a lack of relationship between the variety of resources used and the quality of the proposed solutions. Shih (2019) found that in some cases, although students consulted multiple webpages from different resources, they still could not effectively retrieve the correct information. Olalla-Soler (2018) suggested that the students' information-seeking behaviour was ineffective, so the quality of their translations could not have benefited from an increase in the number of resource types used. We agree with this explanation and believe that the results of our study also stem from the ineffectiveness of student translators' consultation behaviour. Although they consulted more web pages or more resource types, they did not find the information that would have been most useful for solving their translation problems. Therefore, the increase in consultation complexity did not result in an improvement in translation acceptability. This finding reveals that trainee translators are not skilful in using online resources, especially in how to select the most helpful and efficient resource types. It suggests that online consultation skills for translation should be given more weight in translation curriculum design, for example 'asking students to translate a text with limited queries for each translation problem, or to translate a text while being screen-recorded

and then view their translation process recordings and write a report on their information-seeking processes' (Olalla-Soler 2018, p.1313).

6. Conclusion

This study explored the effect of translation difficulty on consultation, the correlation between types of translation problem and consultation, and the effect of consultation on translation acceptability. Our findings can be summarised as follows: First, as translation difficulty increases, consultation represents a larger proportion of the entire translation process and becomes more complex. However, a higher difficulty level does not lead to an increase in consultation difficulty. Second, as the type of translation problem changes, the amount of cognitive effort devoted to consultation and information evaluation behaviour changes accordingly. With an increase in the need for more complex information rather than the explanation of words, participants devote more cognitive effort to extra-linguistic information. Third, an increase in consultation length leads to an increase in translation acceptability; however, no significant correlation was found between consultation complexity and translation acceptability.

The results of this study will contribute to the development of the investigation of consultation behaviour in translation and provide suggestions for translator training. However, we are aware that some limitations exist in this study, including participant type limitations and the use of a single text type and domain. Our future study will include data from professional translators, and add source texts from different domains. In addition, scanpath, which indicates the processing patterns of consultation in translation, will also be included in our analysis and discussion.

References

Angelone, E. (2010). Uncertainty, uncertainty management and metacognitive problem solving in the translation task. In G. M. Shreve & E. Angelone (Eds.), *Translation and cognition* (pp. 17–40). Amsterdam and Philadelphia: John Benjamins.

Daems, J., Carl, M., Vandepitte, S., Hartsuiker, R., & Macken, L. (2016). The effectiveness of consulting external resources during translation and post-editing of general text types. In M. Carl, S. Bangalore, & M. Schaeffer (Eds.), *New directions in empirical translation process research* (pp. 111–133). London and New York: Springer.

Ellis, D. (1989). A behavioural approach to information retrieval design. *Journal of Documentation*, 45(3), 171-212.

Enríquez Raído, V. (2011). Developing web searching skills in translator training. *Revista Electrónica de Didáctica de La Traducción e Interpretación*, (6), 60–80. http://hdl.handle.net/2292/15717.

Enríquez Raído, V. (2014). *Translation and web searching*. New York and London: Routledge. https://doi.org/10.4324/9780203798034

Fulford, H., & Zafra, J. G. (2004). The uptake of online tools and web-based language resources by freelance translators: Implications for translator training, professional development, and research. In E. Y. Rodrigo (Ed.), *Proceedings of the second international workshop on language resources for translation work, research and training* (pp. 50–57). Geneva: The University of Geneva.

Gough, J. (2016). *The patterns of interaction between professional translators and online resources.* (Unpublished doctoral dissertation). University of Surrey.

Granka, L. A., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. Proceedings of the 27th Annual International ACM SIGIR Conference on research and development in information retrieval. (pp. 478–479). https://doi.org/10.1145/1008992.1009079.

Hvelplund, K. T. (2014). Eye tracking and the translation process: reflections on the analysis and interpretation of eye-tracking data. In R. M. Martín (Ed.), *MonTI Special Issue 1: Minding translation | Con la traducción en mente* (pp. 201–223). http://dx.doi.org/10.6035/MonTI.2014.ne1.6

Hvelplund, K. T. (2017). Translators' use of digital resources during translation. *Journal of Language and Communication in Business*, (56), 71–87. https://doi.org/10.7146/hjlcb.v0i56.97205

Hvelplund, K. T. (2019). Digital resources in the translation process–attention, cognitive effort and processing flow. *Perspectives*, 27(4), 510–524.

Jensen, K. T. H. (2009). Indicators of text complexity. In S. Göpferich, A. L. Jakobsen & I. M. Mees (Eds.), *Behind the mind: Methods, models and results in translation process research* (pp. 61–80). Samfundslitteratur.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspectives. *Journal of the American Society for Information Science,* 42(5), 361-371.

Kuznik, A., & Olalla-Soler, C. (2018). Results of pacte group's experimental research on translation competence acquisition. The acquisition of the instrumental sub-competence. *Across Languages and Cultures*, 19(1), 19–51. https://doi.org/10.1556/084.2018.19.1.2

Liu, Y., Zheng, B., & Zhou, H. (2019). Measuring the difficulty of text translation. *Target*, 31(1), 125–149. https://doi.org/10.1075/target.18036.zhe

Miranda, S. V., & Tarapanoff, K. M. A. (2008). Information needs and information competencies: A case study of the off-site supervision of financial institutions in Brazil. *Information Research*, 13(2), 1–24.

Navarro-Prieto, R., Scaife, M., & Rogers, Y. (1999). Cognitive Strategies in Web Searching. *Proceedings on the 5th Conference on Human Factors & the Web*, 1–12. Retrieved from http://zing.ncsl.nist.gov/hfweb/proceedings/navarro-prieto/

O'Hagan, M. (2012). Impact of new technologies on translation studies. In C. Millán & F. Bartrina (Eds.), *The Routledge Handbook of Translation Studies* (pp. 503–518). London and New York: Routledge.

Olalla-Soler, C. (2018). Using electronic information resources to solve cultural translation problems. Differences between students and professional translators. *Journal of Documentation*, 74(6), 1293–1317.

PACTE. (2009). Results of the validation of the PACTE translation competence model: Acceptability and decision making. *Across Languages and Cultures*, 10(2), 207–230. https://doi.org/10.1556/acr.10.2009.2.3

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

Salojärvi, J., Kojo, I., Simola, J., & Kaski, S. (2003). Can relevance be inferred from eye movements in information retrieval. *Proceedings of the 4th Workshop on Self-Organizing Maps* (WSOM 2003), 261–266. https://doi.org/10.1145/1460096.1460120

Shih, C. Y. (2019). A quest for web search optimisation: An evidence-based approach to trainee translators' behaviour. *Perspectives*, 27(6), 908–923. https://doi.org/10.1080/0907676X.2019.1579847

Shneiderman, B. (1997). Designing information-abundant web sites: Issues and recommendations. *International Journal of Human-Computer Studies*, 47(1), 5–29. https://doi.org/10.1006/ijhc.1997.0127

Sycz-Opoń, J. (2019). Information-seeking behaviour of translation students at the University of Silesia during legal translation–an empirical investigation. *Interpreter and Translator Trainer*, 13(2), 152–176. https://doi.org/10.1080/1750399X.2019.1565076

Vakkari, P. (1999). Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval. *Information Processing & Management*, 35(6), 819–837. https://doi.org/10.1016/s0306-4573(99)00028-x

White, M. D., & Iivonen, M. (2001). Questions as a factor in Web search strategy. *Information Processing and Management*, 37(5), 721–740. https://doi.org/10.1016/S0306-4573(00)00043-1

White, M. D., Matteson, M., & Abels, E. G. (2008). Beyond dictionaries: Understanding information behavior of professional translators. *Journal of Documentation*, 64(4), 576–601. https://doi.org/10.1108/00220410810884084

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249–270. https://doi.org/10.1108/eb026404

Wilson, T. D. (2000). Human information behavior. *Informing Science: The International Journal of an Emerging Transdiscipline*, 3(2), 049–056. https://doi.org/10.28945/576

Appendix I

Warm-up text

Source: *Sciencedaily*

Plant pollens vary in quality as food sources for bees, and pollen from the sunflower family is known to have some unpleasant qualities. Bees fed exclusively sunflower pollen often develop poorly, slowly, or not at all. Yet many bee species collect pollen exclusively from this family; in fact, specialisation on sunflower pollen has evolved multiple times in bees.
(Number of words: 58)

Text A

Source: *New Scientist*

There was a time when we thought humans were special in so many ways. Now we know better. We are not the only species that feels emotions, empathises with others or abides by a moral code. Neither are we the only ones with personalities, cultures and the ability to design and use tools. Yet we have steadfastly clung to the notion that one attribute, at least, makes us unique: we alone have the capacity for language. Alas, it turns out we are not so special in this respect either. Key to the revolutionary reassessment of our talent for communication is the way we think about language itself.
(Number of words: 107)

Text B

Source: *Coral Reef and Global Climate Change*

Coral reefs have the highest biodiversity of any marine ecosystem, and they provide important ecosystem services and direct economic benefits to large and growing human populations in coastal zones. Although the natural habitat of coral reefs can be a stressful environment, recent global increases in reef ecosystem degradation and mortality suggest that the rate and nature of recent environmental changes often exceed the adaptive capacity of coral reefs. This crisis is almost certainly the result of interactions between multiple stresses. These include increased nutrient and sediment loading, direct destruction, contamination, overharvesting, disease and predation. Rising ocean temperatures have been implicated in chronic stress, disease epidemics, mass coral bleaching episodes and reduced calcification.
(Number of words: 112)

Appendix II

Questionnaire about educational background and familiarity with background knowledge of source texts

1. Age:
2. Gender:
3. Highest Education Level:

4. Please indicate your first language and second language:
   First language: Second language:

5. At what age did you start to learn your second language?

6. Please indicate your highest IETLS score.
   Overall: Listening: Reading: Writing: Speaking:

7. How many years of translation training do you have?

8. Please indicate your familiarity with the background knowledge of Text A.
   1 2 3 4 5 6 7 8 9 10
   Low High

9. Please indicate your familiarity with the background knowledge of Text B.
   1 2 3 4 5 6 7 8 9 10
   Low High

Appendix III

Permutations of Acceptability (cf. PACTE 2009, p. 218)

Meaning	Function	Language	Assessment	Numerical value
A	A	A		
A	A	SA		
A	SA	A	**A**	**1**
A	SA	SA		
SA	A	A		
A	A	NA		
A	SA	NA		
A	NA	A		
A	NA	SA	**SA**	**0.5**
SA	SA	A		
SA	SA	SA		
SA	A	SA		
A	NA	NA		
SA	SA	NA		
SA	A	NA		
SA	NA	A		
SA	NA	SA		
SA	NA	NA		
NA	A	A	**NA**	**0**
NA	SA	A		
NA	SA	SA		
NA	SA	NA		
NA	A	SA		
NA	A	NA		
NA	NA	A		
NA	NA	SA		