The impact of verbal working memory on written translation: Empirical evidence and an initial model

Jie Li
Leipzig University, Germany and Beijing Foreign Studies University, China
leonie_ljie@outlook.com
https://orcid.org/0000-0001-9740-7808

Abstract

Working memory refers to a metaphorical mental workspace involved in the temporary storage and processing of information which has been considered a key construct in Cognitive Translation Studies. This study investigates whether verbal working memory in the translator’s first and/or second language is related to written translation speed and quality. Through three selection tests, 33 student translators in the second or third year of their Master’s in Translation and Interpreting at nine Chinese universities were selected to participate. In the first test, processing and storage capacities of verbal working memory in their first and second language were assessed separately by means of a dual-task paradigm implemented in the e-Prime software suite. In the second study, participants translated a technical text from their second language to their first language. Their behavioural data were collected with a keylogger and the quality of their translations was evaluated by four expert translators. The correlation analysis suggests that the processing function of verbal working memory in the second language does affect logged translation speed; and that the processing function of verbal working memory in the first language has an effect on translation quality. Furthermore, through regression analysis, an initial model was established, which shows the precise structure of the impact of bilingual working memory on written translation. These findings have significant implications for translation skills training. It appears that additional verbal working memory training for both first and second languages would enhance the students’ translation skills.

Keywords: cognitive translation studies; cognitive translation effort; working memory and translation; verbal working memory; translation process

1. Introduction

“Our memory is a more perfect world than the universe: it gives back life to those who no longer exist.”

Guy De Maupassant.

In order to solve an arithmetic problem such as \((4 \times 2) - (5 \times 3)\) in our head, without writing it on paper, or to read a text and then summarize it without taking notes, or to follow what the professor says until the end in a hard class, we need a kind of memory that helps us to retain ideas and thoughts as we work on them. This type of memory is called working memory (WM). It is usually divided into three types: numerical, spatial–figurative and verbal (Süß et al., 2002) – the last of these is the focus of this article.

WM is “a multi-component system responsible for active maintenance of information in the face of ongoing processing and/or distraction” (Conway et al., 2005, p. 770). As a cognitive system (Baddeley & Hitch, 1974), it metaphorically refers to a mental workspace or a piece of mental jotting paper, with two separate functions: first, a temporary storage function analogous to short-term memory and, second, an information-processing function that controls the flow of information between the short-term storage systems and long-term memory (LTM) (e.g., Baddeley, 1986, 2003; Cowan, 1988; Juffs & Harrington, 2011).

Four main models of the cognitive mechanism of these two subcomponents have been proposed (e.g., Li, 2019; Miyake et al., 2001; Oberauer, 2009; Sun et al., 2012):

- The resource-sharing model (e.g., Case et al., 1982; Daneman & Carpenter, 1980; Just & Carpenter, 1992) suggests that both processing and storage require attention, which is a limited resource. It leads to a trade-off hypothesis of resource availability within the subcomponents.
- Following the task-switching model by Towse et al. (2001), WM performance may be associated with difficulty in task-switching from storing information to processing it.
- The time-based resource-sharing model (e.g., Barrouillet et al., 2004; Towse et al., 2001) assumes that attention switches between processing and storing and that, as a consequence, this task-switching delays recall while preventing rehearsal. This means that memory traces decay over time.
- The interference-based model by Saito and Miyake (2004) and Oberauer and Lewandowsky (2008) finds that the larger the number of items to which attention has to be paid, the larger the reduction in the speed and accuracy of recall.

One point on which these models disagree is whether WM, in particular verbal WM, is language-specific or language-independent. This led to the formulation of the following research question, which this article attempts to answer: Will a translator exhibit consistent (relative) verbal WM performance from language to language or will their performance differ substantially across languages? On the one hand, Marton and Schwartz (2003), Marton et al. (2006) and Maurits et al. (2006) support the view that verbal WM performance depends on
language proficiency because of differences in language structure, that is, in grammar, morphology and syntax, etc. Proponents of this view would, for instance, expect that a translator with excellent proficiency in English but only mediocre proficiency in Japanese would have correspondingly high WM performance in English and middling performance in Japanese. On the other hand, Harrington and Sawyer (1992), Osaka et al. (1993), Miyake and Friedman (1998) and Swanson et al. (2004) argue that verbal WM performance is immune to language differences. They would expect that the translator would have similar WM performance across English and Japanese, despite differences in proficiency (note that the overall translation quality would still differ). Sun et al. (2012) have a more nuanced view, concluding that verbal WM-processing is cross-linguistic, whereas verbal WM storage is language-specific.

In the cognitive approach to Translation Studies, verbal WM has been considered as a key construct. Many scholars have explored the relationship between verbal WM and interpreting (e.g., Christoffels et al., 2003, 2006; Cowan, 2000; Danks et al., 1997; Gile, 1997; Lederer, 1978; Liu, 2003; Mizuno, 2005; Moser-Mercer, 2000; Timarrová, 2008; Zhang, 2011). They assume that the interpreter perceives the source-language (SL) expression, stores it and translates it into the target language (TL) within a short period of time (e.g., Bao, 1998; Gile, 1997; Zhang, 2011). Theoretical paradigms specific to the mental processes involved in interpreting have been put forward, such as by Darò and Fabbro’s (1994) model of WM in simultaneous interpreting, Gile’s (1997) effort models and Mizuno’s (2005) embedded process model for simultaneous interpreting.

Compared to the above research in interpreting, research into the effects of verbal WM in written translation is underdeveloped. It has generally been assumed that, whereas the interpreter heard it only once, translators could reread the original text at will and therefore the role of verbal WM was of no importance in written translation. But other studies have cast doubt on this view.

Seleskovitch (1978) developed the Interpretive Theory of Translation (ITT) and proposed that the SL text should not be translated word for word into the TL, but rather that “the interpreter should separate meaning from wording” (p. 36). Lederer (1994) assumed that cognitive memory in translation is also necessary, because in the process of deverbalization not words but “units of meaning” (p. 23) are stored in the mind (Seleskovitch 1978, in contrast to Kintsch, 1998). In addition, the translator combines the temporarily stored information from the original text with those retained in LTM, such as vocabulary or grammar in both languages. Regarding the resource-sharing model of WM (e.g., Case et al., 1982; Daneman & Carpenter, 1980; Just & Carpenter, 1992), in written translation the translator’s attention must be switched between the understanding of the SL text and the representation of the TL text, and also between storing information and processing it. Therefore, it stands to reason that WM is involved in written translation, as has been posited by Liu (2015).

Many studies have investigated verbal WM. Some focused on the dynamic relationship between the storage and processing of verbal WM (e.g., Baddeley, 1986; Cowan, 1988; Juffs & Harrington, 2011; Oberauer, 2009; Sun et al., 2012), whereas others have focused on the

Language-(in)dependence of verbal WM performance (e.g., Harrington & Sawyer, 1992; Marton & Schwartz, 2003; Marton et al., 2006; Miyake & Friedman, 1998). However, many questions have been left unanswered:

- Do the storage span and the processing capacity of verbal WM have an interference effect or a facilitatory effect on each other?
- Is the capacity of verbal WM language-specific or language independent? In other words, is there a difference in verbal WM capacity between the translator’s native tongue (L1) and the target foreign language (L2)?
- Is verbal WM in L1 and/or L2 related to written translation speed and quality?
- If there is such a relationship, what is its exact nature?

2. Research methods

Three selection tests (Appendices I–III) – including a physics quiz on electricity and magnetism, a questionnaire about language proficiency, computer operation skills, etc. and a standard progressive matrices (SPM) IQ test (Raven, 1941) – were created in order to ensure that the participants have similar abilities in areas which would affect their translations. The PACTE translation competence model (2003, p. 60, see Figure 1) was used to assess which areas are relevant in this respect and which should be controlled for with selection tests. It contains five subcompetences (i.e., bilingual, extra-linguistic, instrumental, strategic and knowledge of translation) and psychophysiological components.

Figure 1. Translation competence model (from PACTE, 2003, p. 60)

Following the guideline proposed by Roscoe (1975), a sample size of at least 30 is deemed necessary in order to ensure the tests’ results are accurate and sound. Such a sample size is typical for studies of this nature – for example, those of Rothe-Neves (2003) and Wang and Xu (2009). Initially, 35 students in their second or third year of a Master’s degree in Translation

and Interpreting from nine Chinese universities were invited to complete the three online selection tests. Through these three tests, their bilingual competence, instrumental competence, knowledge about translation, IQ-level and the indicators that reflect participants’ physical functions such as their left- or right-handedness and vision, could be measured.

The purpose of the selection tests was to keep extraneous variables controlled (see Figure 2), such as bilingual competence of the participants, that may interfere with the results of the study designed to investigate the effect of verbal WM on written translation.

![Figure 2. Overview of studies](image)

After taking the selection tests, 33 of the students were finally chosen to participate in the following studies (concerning the others: one was not familiar with the use of online dictionaries and search engines and the other obtained a lower score in the electricity quiz).

- Thirty-one of the participants were female and two were male. The test group was overwhelmingly female because of the high ratio of female to male students in Chinese–German programmes. The average age was 24 years. All of them were right-handed and had normal vision.
- All of them had Chinese as L1 and their Master’s major was translation between German and Chinese. All of them had either a test of *Deutsch als Fremdsprache* (German as a foreign language) certificate (with more than 18 points out of 20), a test of *Prüfung für das Germanistik-Hauptstudium* (College entrance examination for German Studies) (with a score of more than 70 out of 100) or a China Accreditation Test for Translators and Interpreters (CATTI) with level II out of level IV, and all of
them had spent less than one year in German-speaking countries.

- They were familiar with the use of online dictionaries and search engines and had experience translating on Windows laptops.
- Their background knowledge of electricity and magnetism was similar (their scores were between 20 and 50 out of 100), which is the theme of the translation text in the following studies.
- Their IQ-level scores in the Standard Progressive Matrices IQ test were between 44 to 57 out of 60.

They did not know the purpose of the study until it was completed. This measure served to ensure that the feedback and the data collected were genuine.

### 2.1. Study I: Verbal working memory

The first study investigated the processing and storage functions of verbal WM in L1 and L2. It was programmed and implemented in the e-Prime software suite, which contained a test-set covering all language and function combinations (see Figure 3). Both the processing and the storage functions of verbal WM in L1 and L2 were assessed separately via a set of tests based on a verbal and visual dual-task paradigm (e.g., Bayliss et al., 2003; Shah & Miyake, 1996) in which the participants received both a verbal and a visual stimulus in the same trial. This means that the verbal WM capacity was measured via the retention interval of a visuospatial WM task.

Tests 1 and 2 measured the verbal WM storage span. In each round, each participant received a visuospatial reasoning task which required the participants to identify whether various graphics were symmetrical or asymmetrical. Then a verbal span task was presented which required the participants to recall a list of Chinese words (in test 1) or German words (in test 2). In this manner, visuospatial reasoning tasks and verbal span tasks appeared alternately.

The experimental procedure details are as follows:

- First, a red cross appeared as a fixation point in the centre of the screen for 1.3 s.
- Next, one of the 27 words (in test 1, a Chinese word; in test 2, a German word) appeared randomly on the screen for 1.8 s. During this time, the participant had to read the presented word aloud and memorize it without writing it down.
- Then one of the 27 pictures appeared randomly on the screen. The picture remained on screen until the participant responded by pressing one of the two buttons on the keyboard. Here, the participants had to indicate as quickly as possible whether each picture presented was symmetrical (by pressing the “j” button) or asymmetrical (by pressing the “f” button).
- Then another round started.

At the end of each block, the participants were required to recall all the words within that block, in order. If the answer was correct, then the participant continued to the next block; if it was not correct, the test terminated. Those participants who made it to the sixth block were tasked with recalling at most seven German or Chinese words.

Tests 3 and 4 were tests of the verbal WM processing capacity. A visuospatial span task which requires participants to recall the spatial positions of a set of graphics, and a verbal reasoning task which requires them to answer whether various Chinese sentences (in test 3) or German sentences (in test 4) are grammatically correct, are embodied in them and appear one after the other. The test process proceeded as follows:

- At the beginning, a red fixation point appeared for 1.3 s to cue the participants that they needed to pay attention to the following task.
- Then, one of the 14 sentences (Chinese in test 3 and German in test 4) appeared randomly on the screen. At this point, the participant had to indicate as soon as possible whether the presented sentence was grammatically correct (by pressing the “j” button) or incorrect (by pressing the “f” button). After they had keyed in their response, the sentence disappeared.
- Then the participant saw one of the 14 pictures of a grid consisting of 4 × 4 squares, with one square randomly painted red. The square remained visible for 2.3 s. The participant, during those 2.3 s, needed to memorize the position of the red cube without taking notes.

Each test in the study consisted of multiple blocks (tests 1 and 2 had from one to six, depending on the participant’s performance; tests 3 and 4 had four blocks), each block contained multiple rounds (two for the first block, with each subsequent block having one more round than the previous one). At the end of each block, the participant had to mark all the positions of the presented red squares in this round on the test paper provided. In total, the test could last for four blocks. That meant the participant would need to recall at most five spatial positions and judge five sentences in one round.

**Figure 3. Test-set in study I**

### 2.2. Study II: Working memory and translation performance

In the second study, the participants were required to translate a technical text from L2 to L1. The source text (ST) was an excerpt from a German patent (Smithies & Hale, 1968, p. 1) on electricity and magnetism. It consisted of five paragraphs and contained a total of 196 German words. The frequency ratio of technical terms in the text was 14.29%.

Translog II was used to obtain data relevant to the translation process. This keylogging software divides the screen horizontally into an ST window (top) and a target text (TT) window (bottom) where translators type their translations. During translation, the inputs from the keyboard and the mouse are automatically recorded, allowing researchers to obtain data reflecting the participant’s translation process and to measure their logged translation speed.

Four expert translators evaluated the quality of the translations using the assessment model of He (2015, see Figure 4), which is based on Saaty’s (1980) analytic hierarchy process for complex decision-making. It combines qualitative and quantitative analysis steps, and also offers a systematic and hierarchical analysis (He, 2015, p. 208). The approach is very adequate for assessing translations of technical or scientific texts (Zhou et al., 1990) and has been used extensively (He & Zhang, 2012).

Figure 4. Translation quality assessment model of He (2015)

The translations are evaluated at three main levels: purpose (A), criteria (B), and indicator (C).

Manafi (2004) defines the purpose of translation as “accurate translation … which is the reproduction of the message of the ST [source text]” (p. 34). The lexical–semantic dimension (B₁) contains three sub-indicators, lexicon (C₁.1), syntax (C₁.2) and pragmatics (C₁.3). The textual dimension (B₂) contains three more indicators: text style (C₂.1), text structure (C₂.2) and text function (C₂.3). The cognitive dimension (B₃) also has three indicators: flow of ideas (are the transitions between ideas fluent? C₃.1), metaphor (do metaphors convey the same effect as in the ST? C₃.2) and aesthetic (does the translation retain aesthetic qualities? C₃.3).

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Indicator</th>
<th>score (0–5 points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lexical–semantic dimension</td>
<td>Lexicon</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Syntax</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pragmatics</td>
<td></td>
</tr>
<tr>
<td>textual dimension</td>
<td>text style</td>
<td></td>
</tr>
<tr>
<td></td>
<td>text structure</td>
<td></td>
</tr>
<tr>
<td></td>
<td>text function</td>
<td></td>
</tr>
<tr>
<td>cognitive dimension</td>
<td>the flow of ideas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>metaphoric representation of ideas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>aesthetic representation of ideas</td>
<td></td>
</tr>
<tr>
<td><strong>Scale:</strong></td>
<td>0–1.0, unsatisfactory; 1.1–2.0, poor; 2.1–3.0, adequate; 3.1–4.0, satisfactory; 4.1–5.0, good;</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4 shows the weighting coefficients of the lexical–semantic (B1), textual (B2) and cognitive (B3) criteria, which were 0.4, 0.3 and 0.3 respectively. At the indicator level, the weighting coefficient for all nine sub-indicators was 0.33.

The expert translators were tasked with evaluating the translated text from the participants and giving a score (between 0 and 5) to the nine sub-indicators. If $N_1$ represents the score for the lexical—semantic dimension, $N_2$ the score for the textual dimension etc. and, similarly, $M_{1,1}$ the score for the lexicon, $M_{1,2}$ the score for the syntax and so forth for the other indicators, then the score for criterion $N_i$ where $i \in \{1, 2, 3, 4\}$ can be calculated with the following formula:

$$ N_i = \sum_{j=1}^{3} C_{i,j} M_{i,j} = C_{1,1} C_{1,2} C_{1,3} \left( \frac{M_{1,1}}{M_{1,3}} \right) $$

For example, one expert translator might evaluate the translation of one participant by marking the nine sub-indicators as follows:

**Table 2. An example of evaluation**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Criterion</th>
<th>Indicators</th>
<th>Notes for each indicator</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>lexical–semantic dimension</td>
<td>$M_{1,1}$</td>
<td>4.5; 4; 4</td>
<td>4.1745</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_1$</td>
<td>4.125</td>
<td></td>
</tr>
<tr>
<td></td>
<td>textual dimension</td>
<td>$M_{2,1}$</td>
<td>4.5; 4.5; 4.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_2$</td>
<td>4.455</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cognitive dimension</td>
<td>$M_{3,1}$</td>
<td>3.5; 4; 4.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$N_3$</td>
<td>3.96</td>
<td></td>
</tr>
</tbody>
</table>

According to the above formula, the scores for each criterion would be calculated as follows:

$$ N_1 = \sum_{j=1}^{3} C_{1,j} M_{1,j} = 4.5 \times 0.33 + 4 \times 0.33 + 4 \times 0.33 = 4.125 $$

$$ N_2 = \sum_{j=1}^{2} C_{2,j} M_{2,j} = 4.5 \times 0.33 + 4.5 \times 0.33 + 4.5 \times 0.33 = 4.455 $$

$$ N_3 = \sum_{j=1}^{3} C_{2,j} M_{2,j} = 3.5 \times 0.33 + 4 \times 0.33 + 4.5 \times 0.33 = 3.960 $$

Then, the total score for this translation \( (Q) \) would be calculated according to the following formula:

\[
Q = (B_1, B_2, B_3) \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}
\]

Our example expert translator would give this translation a score of 4.1745:

\[
= (B_1, B_2, B_3) \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}
= (0.4, 0.3, 0.3) \begin{pmatrix} 4.125 \\ 4.455 \\ 3.960 \end{pmatrix}
\]

Finally, the final score for Q is the average of the scores given by the four expert translators.

3. Results and discussion

This study is based on a four-year PhD project conducted at Leipzig University. The results of the first study were presented and discussed in Li (2019). According to the findings, the performance of verbal WM is language-specific and its subcomponents, that is, processing and storage, have a positive impact on each other (Li, 2019).

3.1 Storage span for verbal working memory

In tests 1 and 2, the response-time and correctness of the visual judgement performance were automatically recorded by e-Prime, whereas we manually recorded the correctness of the verbal-recall task and calculated the specific WM-storage span of the participants according to the following formula (let \( n \) represent the block number; \( n + 1 \), the latest round in this block; and \( a \) the error counters (including recall mistakes and forgotten items) in this block):

\[
\text{storage span} = (n + 1) - a \times \frac{1}{(n + 1)}
\]

The marking scale is a 7-point grading scale. For example, if a participant (e.g., in test 1, fifth block) correctly recalls four Chinese words, forgets one word and recalls one word incorrectly, their WM-storage span in L1 equals 5.67 points \((6 - 2 \times (1/6))\).

The verbal-recall performance both in L1 (Chinese) and in L2 (German) of all the 33 participants was calculated using the formula given above. The results are given in Table 3.

It had been assumed that the storage span for verbal WM in L1 should be greater than in L2. This was indeed the case for each participant and was reflected in the average results where the average verbal WM storage span in L1 (6.462) was greater than in L2 (4.173). From these results, it can be concluded that there is a language-specific difference in the verbal WM storage span among the participants.

**Table 3. Verbal WM storage span in both languages from 33 participants**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Verbal WM storage span</th>
<th>Participant</th>
<th>Verbal WM storage span</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L1</td>
<td></td>
<td>L2</td>
</tr>
<tr>
<td>1</td>
<td>7.00</td>
<td>18</td>
<td>7.00</td>
</tr>
<tr>
<td>2</td>
<td>6.57</td>
<td>19</td>
<td>6.86</td>
</tr>
<tr>
<td>3</td>
<td>6.71</td>
<td>20</td>
<td>5.50</td>
</tr>
<tr>
<td>4</td>
<td>6.86</td>
<td>21</td>
<td>7.00</td>
</tr>
<tr>
<td>5</td>
<td>6.86</td>
<td>22</td>
<td>6.86</td>
</tr>
<tr>
<td>6</td>
<td>6.71</td>
<td>23</td>
<td>7.00</td>
</tr>
<tr>
<td>7</td>
<td>6.86</td>
<td>24</td>
<td>6.43</td>
</tr>
<tr>
<td>8</td>
<td>6.86</td>
<td>25</td>
<td>6.71</td>
</tr>
<tr>
<td>9</td>
<td>5.67</td>
<td>26</td>
<td>5.83</td>
</tr>
<tr>
<td>10</td>
<td>5.83</td>
<td>27</td>
<td>6.86</td>
</tr>
<tr>
<td>11</td>
<td>5.83</td>
<td>28</td>
<td>5.67</td>
</tr>
<tr>
<td>12</td>
<td>7.00</td>
<td>29</td>
<td>7.00</td>
</tr>
<tr>
<td>13</td>
<td>6.86</td>
<td>30</td>
<td>5.17</td>
</tr>
<tr>
<td>14</td>
<td>7.00</td>
<td>31</td>
<td>5.83</td>
</tr>
<tr>
<td>15</td>
<td>6.71</td>
<td>32</td>
<td>6.71</td>
</tr>
<tr>
<td>16</td>
<td>5.83</td>
<td>33</td>
<td>6.86</td>
</tr>
<tr>
<td>17</td>
<td>4.80</td>
<td>Mean</td>
<td>6.462</td>
</tr>
<tr>
<td></td>
<td>3.50</td>
<td></td>
<td>4.173</td>
</tr>
</tbody>
</table>

**3.2 Verbal working memory – processing capacity**

In tests 3 and 4, the response-time and correctness of the sentence judgement task was automatically recorded with the help of e-Prime. Then, a standard score (Z-score) was introduced in order to facilitate the comparison between the participants. It was calculated with the following formula: let $x$ be the participant’s score; $\bar{x}$ the mean; and $S$ the standard deviation of the sample:

$$Z = \frac{x - \bar{x}}{S}$$

In order to assess the verbal processing performance in both languages, the Z-score for response-time ($Z_{rt}$) and that for response accuracy ($Z_a$) were calculated separately. Based on these two scores, we calculated a combined Z-score for verbal WM processing-performance ($Z_t$) using the following formula:

$$Z_t = Z_a - Z_{rt}$$

For example, a participant needed on average 4872.64 ms to judge one Chinese sentence in test 3, and she judged 12 of the 14 sentences correctly. Her Z-score for response-time ($Z_{rt}$) was $-0.73$ (i.e. $[4872.64 - 6099.20]/1689.30$) and their Z-score for response accuracy ($Z_a$) was 2.08 (i.e. $[0.86 - 0.63]/0.11 = 2.08$). According to their behavioural data (i.e., response-time and accuracy), they got a score of 2.81 points (i.e. $2.08 - (-0.73)$) for their verbal WM processing capacity in L1. The scores for verbal WM processing capacity in both languages for all 33 participants were calculated. They are shown in Table 4 below.

**Table 4. Verbal processing performance in both languages from 33 participants**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Verbal processing performance</th>
<th>Verbal processing performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L1</td>
<td>L2</td>
</tr>
<tr>
<td>1</td>
<td>2.81</td>
<td>1.79</td>
</tr>
<tr>
<td>2</td>
<td>2.60</td>
<td>0.87</td>
</tr>
<tr>
<td>3</td>
<td>0.72</td>
<td>0.52</td>
</tr>
<tr>
<td>4</td>
<td>-0.22</td>
<td>-0.79</td>
</tr>
<tr>
<td>5</td>
<td>2.67</td>
<td>1.03</td>
</tr>
<tr>
<td>6</td>
<td>0.55</td>
<td>1.36</td>
</tr>
<tr>
<td>7</td>
<td>0.78</td>
<td>1.99</td>
</tr>
<tr>
<td>8</td>
<td>0.95</td>
<td>1.92</td>
</tr>
<tr>
<td>9</td>
<td>-0.02</td>
<td>1.63</td>
</tr>
<tr>
<td>10</td>
<td>1.64</td>
<td>-0.54</td>
</tr>
<tr>
<td>11</td>
<td>-1.92</td>
<td>-0.72</td>
</tr>
<tr>
<td>12</td>
<td>2.02</td>
<td>-0.30</td>
</tr>
<tr>
<td>13</td>
<td>-1.02</td>
<td>0.16</td>
</tr>
<tr>
<td>14</td>
<td>1.23</td>
<td>-4.79</td>
</tr>
<tr>
<td>15</td>
<td>-2.69</td>
<td>0.55</td>
</tr>
<tr>
<td>16</td>
<td>-0.79</td>
<td>0.11</td>
</tr>
<tr>
<td>17</td>
<td>0.80</td>
<td>-1.04</td>
</tr>
</tbody>
</table>

Whether there is a correlation between the verbal WM processing capacity in L1 and L2 is investigated in the analysis in Section 3.5.

**3.3. Logged translation speed**

By using Translog II, the logged translation speed was automatically logged (see Table 5). The translation duration of participants was between 1537.08 s and 6592.77 s. The difference is therefore 5055.69 s, which means that the informants had large individual differences. On average, a participant needed 3026.17 s to translate the ST.

### Table 5. Translation efficiency in the study II

<table>
<thead>
<tr>
<th>Participant</th>
<th>Speed</th>
<th>Quality</th>
<th>Participant</th>
<th>Speed</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6592.77</td>
<td>3.68</td>
<td>18</td>
<td>2559.49</td>
<td>3.48</td>
</tr>
<tr>
<td>2</td>
<td>2610.70</td>
<td>3.83</td>
<td>19</td>
<td>2466.53</td>
<td>3.27</td>
</tr>
<tr>
<td>3</td>
<td>2733.05</td>
<td>3.55</td>
<td>20</td>
<td>3085.73</td>
<td>3.20</td>
</tr>
<tr>
<td>4</td>
<td>4807.92</td>
<td>3.94</td>
<td>21</td>
<td>2985.30</td>
<td>3.45</td>
</tr>
<tr>
<td>5</td>
<td>2892.03</td>
<td>3.98</td>
<td>22</td>
<td>6394.48</td>
<td>4.19</td>
</tr>
<tr>
<td>6</td>
<td>2755.17</td>
<td>3.41</td>
<td>23</td>
<td>3053.17</td>
<td>4.16</td>
</tr>
<tr>
<td>7</td>
<td>3396.78</td>
<td>3.46</td>
<td>24</td>
<td>3799.51</td>
<td>3.80</td>
</tr>
<tr>
<td>8</td>
<td>3654.78</td>
<td>3.84</td>
<td>25</td>
<td>3456.03</td>
<td>3.71</td>
</tr>
<tr>
<td>9</td>
<td>2968.24</td>
<td>3.08</td>
<td>26</td>
<td>2898.06</td>
<td>3.62</td>
</tr>
<tr>
<td>10</td>
<td>2969.48</td>
<td>3.80</td>
<td>27</td>
<td>2635.63</td>
<td>2.95</td>
</tr>
<tr>
<td>11</td>
<td>1802.81</td>
<td>3.46</td>
<td>28</td>
<td>1739.50</td>
<td>2.96</td>
</tr>
<tr>
<td>12</td>
<td>2604.80</td>
<td>4.41</td>
<td>29</td>
<td>2562.39</td>
<td>3.16</td>
</tr>
<tr>
<td>13</td>
<td>2425.66</td>
<td>2.91</td>
<td>30</td>
<td>1756.31</td>
<td>3.08</td>
</tr>
<tr>
<td>14</td>
<td>1537.08</td>
<td>3.00</td>
<td>31</td>
<td>2412.95</td>
<td>3.07</td>
</tr>
<tr>
<td>15</td>
<td>2155.98</td>
<td>4.30</td>
<td>32</td>
<td>2941.89</td>
<td>3.34</td>
</tr>
<tr>
<td>16</td>
<td>2630.22</td>
<td>3.19</td>
<td>33</td>
<td>3792.92</td>
<td>3.78</td>
</tr>
<tr>
<td>17</td>
<td>2786.14</td>
<td>3.70</td>
<td><strong>mean</strong></td>
<td><strong>3026.17</strong></td>
<td><strong>3.54</strong></td>
</tr>
</tbody>
</table>

*Note:* Speed measured in seconds; quality expressed in a scale from 0 to 5 points, where 0 = lowest.

### 3.4. Translation quality

As shown in Table 5, 26 participants (78.79%) achieved a satisfactory result of a score between 3 and 4 points. Four participants (12.12%) received a good result with a score of over 4 points. And three participants (9.09%) were rated with a score of less than 3 points, that is, a poor or an unsatisfactory result. The mean translation quality for all participants is 3.54 points and the SD is 0.41. This means that the technical translation skills of the MTI students are overall satisfactory. However, there is room for improvement with more translation training.

### 3.5. Correlation analysis: verbal working memory and written translation

After the collection of all the data from the 33 participants in two studies (namely, the storage span of verbal WM in L1 and in L2, the processing capacity of verbal WM in both languages, logged translation speed and translation quality), a correlation analysis was performed using the SPSS Statistics software package. The Pearson coefficient of correlation “r” indicates the strength and the direction of the linear relationship between two variables. The calculated coefficients suggest the following five significant relationships (see Table 6):

Table 6. Correlations between bilingual verbal WM and translation

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 L1 WM storage</td>
<td>6.46</td>
<td>0.62</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 L2 WM storage</td>
<td>4.13</td>
<td>0.76</td>
<td>0.291</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 L1 WM processing</td>
<td>0.00</td>
<td>1.54</td>
<td>0.355* 0.145</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 L2 WM processing</td>
<td>0.00</td>
<td>1.47</td>
<td>0.257</td>
<td>0.377* 0.176</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 logged transl. speed</td>
<td>3.026.17</td>
<td>1108.16</td>
<td>0.304</td>
<td>0.270</td>
<td>0.290</td>
<td>0.415*</td>
<td>—</td>
</tr>
<tr>
<td>6 translation quality</td>
<td>3.54</td>
<td>0.41</td>
<td>0.302</td>
<td>0.327</td>
<td>0.348*</td>
<td>0.315</td>
<td>0.439*</td>
</tr>
</tbody>
</table>

* p < 0.05 (2-tailed), significant

(1) in the case of L1, the correlation between verbal WM storage span and processing capacity yielded $r = 0.355$, $p = 0.043$, significant;
(2) in the case of L2, the correlation between verbal WM storage span and processing capacity yielded $r = 0.377$, $p = 0.030$, significant;
(3) the correlation between verbal WM processing capacity in L2 and logged translation speed yielded $r = 0.415$, $p = 0.016$, significant;
(4) the correlation between verbal WM processing capacity in L1 and translation quality yielded $r = 0.348$, $p = 0.047$, significant;
(5) the correlation between logged translation speed and translation quality yielded $r = 0.439$, $p = 0.011$, significant.

With the results of this correlation analysis, we can respond to the research questions.

Verbal WM storage span and processing capacity are positively rather than negatively correlated. This can be seen from the significant positive relationships (i.e., $r > 0$ plus a significant $p$-value) between these variables (numbers 1 and 2 above).

There are no significant relationships between verbal WM storage span in L1 and L2 ($r = 0.291$, $p = 0.101$, not significant), between verbal WM storage span in L1 and verbal WM processing capacity in L2 ($r = 0.257$, $p = 0.149$, not significant), between verbal WM processing capacity in L1 and verbal WM storage span in L2 ($r = 0.145$, $p = 0.422$, not significant) or between verbal WM processing capacity in L1 and L2 ($r = 0.176$, $p = 0.326$, not significant). The absence of any significant relationship does not allow us to draw any conclusions about language specificity.

The $p$-value for the correlation coefficient between verbal WM processing capacity in L1 and in L2 ($r = 0.176$) reveals a very broad confidence interval between $-0.209$ and 0.465. In sum,

A significantly larger study (i.e., many more participants) would be needed to answer the question whether verbal WM capacity is language-dependent.

The fifth relationship suggests that there is a significant positive correlation between logged translation speed and quality ($r = 0.439$, $p = 0.011$, significant).

The coefficients corresponding to the third and fourth relationships suggest that there are stronger correlations between verbal WM processing capacity and translation performance: on the one hand, the processing capacity of verbal WM in L1 has a moderate, positive effect on translation quality ($r = 0.348$, $p = 0.047$, significant). On the other hand, the processing capacity of verbal WM in L2 has a moderate, positive effect on the logged translation speed ($r = 0.415$, $p = 0.016$, significant).

3.6. An initial model

After concluding the correlation analysis, we analysed the four types of bilingual WM as independent variables using multiple regression to predict the logged translation speed and quality. Three linear relationships emerged from this analysis (see Table 7):

(1) The impact of L2 verbal WM on logged translation speed.

The value of $R^2$ in the third and fourth linear regression indicates that: the storage span of the verbal WM in L2 explains 14.2% of the processing capacity of the verbal WM in L2; and 17.2% of the variance in the logged translation speed can be predicted from the variable processing capacity. These two coefficients of determination enable the following conclusion to be arrived at: the storage span of the verbal WM in L2 affects the processing capacity of the verbal WM in L2; and the processing capacity of the verbal WM in L2 influences the logged translation speed.

(2) The impact of L1 verbal WM on translation quality

The storage span of the verbal WM in L1 predicts 12.6% of the processing capacity of the verbal WM in L1; and the processing capacity of the verbal WM in L1 determines 11.9% of the translation quality. This strength of association makes possible this conclusion: the storage span of the verbal WM in L1 is associated with the processing capacity of verbal WM in L1; and the processing capacity of the verbal WM in L1 affects the translation quality.

(3) The logged translation speed and quality exert a mutual effect on each other, because the value of $R^2$ (strength of association) was 19.2%.

**Table 7. Linear regression analysis of the relationships between six variables**

<table>
<thead>
<tr>
<th>Independent</th>
<th>Dependent</th>
<th>$R^2$</th>
<th>Linear regression formula</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 storage</td>
<td>L1 processing</td>
<td>0.126</td>
<td>$Y = -5.684 + 0.880x$</td>
<td>0.043*</td>
</tr>
<tr>
<td>L1 processing</td>
<td>translation quality</td>
<td>0.119</td>
<td>$Y = 3.538 + 0.094x$</td>
<td>0.047*</td>
</tr>
<tr>
<td>L2 storage</td>
<td>L2 processing</td>
<td>0.142</td>
<td>$Y = -2.989 + 0.724x$</td>
<td>0.030*</td>
</tr>
<tr>
<td>L2 processing</td>
<td>logged translation speed</td>
<td>0.172</td>
<td>$Y = 3026.357 + 313.341x$</td>
<td>0.016*</td>
</tr>
<tr>
<td>translation quality</td>
<td>logged translation speed</td>
<td>0.192</td>
<td>-</td>
<td>0.011*</td>
</tr>
<tr>
<td>logged translation speed</td>
<td>translation quality</td>
<td>0.192</td>
<td>-</td>
<td>0.011*</td>
</tr>
</tbody>
</table>

With the help of the $R^2$ metric, we established a hierarchical linear model (see Figure 5) which suggests the precise structure of the impact of bilingual WM on written translation.

![Hierarchical linear model](image)

As indicated at the top of the model (see Figure 5), the storage span of the verbal WM in a language serves as a starting parameter that can have a positive influence on the processing capacity of the verbal WM in that language. This means those who have a higher storage span of verbal WM could have a higher processing capacity of verbal WM in the same language.

As seen at the centre left and the left of the model (see Figure 5) the quality of the translation is associated with the processing capacity of the verbal WM in L1. At the centre right and the lower right, the logged translation speed is influenced by the processing capacity of the verbal...

WM in L2. As a result, at least for translations from L2 to L1, those who have a higher processing capacity for verbal WM in L2 could translate much faster because they can understand the information presented in the ST more readily and quickly. And those who have a higher processing capacity for verbal WM in L1 could produce higher-quality translations, because they can more easily find appropriate representations in the TL. From this perspective, the verbal WM in both languages – that is, L1 and L2 – is connected to the logged translation speed and quality.

This model suggests that verbal WM is closely related to translation performance and describes how verbal WM might affect individual aspects of translation. The results of this study suggest that the better the capacity of the WM, the better the translation performance. This confirms that translation not only requires language proficiency in two languages; it also requires, and indeed benefits from, other cognitive resources, in particular the verbal WM in L1 and L2.

4. Conclusions

The research project reported on in this article focused on the possible role of verbal WM in translation. Most prior research efforts focused on interpreting, and there was a need to investigate the impact of verbal WM on translation performance and to describe the exact nature of any interactions between WM components. This project consisted of two studies. The first study investigated the interaction between storage span and processing capacity and the language-specificity of the verbal WM; in this study, the capacities of verbal WM in L1 and L2 were investigated. In the second study, the participants translated a technical text from their L2 into their L1. Their logged translation speed was measured with Translog II and the quality of their translations was evaluated.

Correlation analyses led to the following conclusions:

1. For each language, there is a positive correlation between the storage span and the processing performance of the verbal WM in that language.
2. Verbal WM performance, both in the storage span and the processing capacity, is language-dependent. The participants performed better in their L1 (both for storage span and processing capacity) than in their L2.
3. When translating from L2 into L1, the processing capacity of the verbal WM in L2 affects the logged translation speed; and the processing capacity of the verbal WM in L1 has an effect on the translation quality.

A hierarchical linear model was generated through regression analysis which hints at the precise structure of the impact of bilingual WM on translating: first, the storage span of the verbal WM in each language is a starting parameter that positively influences the processing capacity of the verbal WM in that language. Then, the processing capacity of the verbal WM in the L1 affects the translation quality, whereas the processing capacity of the verbal WM in the L2 influences the logged translation speed.

These results differ somewhat from those of Rothe-Neves (2003) on the influences of WM on some formal aspects of translation performance. This may be related to sample size (Rothe-Neves had only six participants in each group, 12 participants in total) and to the holistic approach to WM (Rothe-Neves did not divide WM into two subcomponents, storage and processing, but took it as a whole). In our results, the storage of WM has a very low direct relationship with translation performance, but the processing of WM has a moderately significant direct relationship with translation performance.

Overall, this study supports the multifunctional structure and language specificity of bilingual verbal WM and its effects on logged translation speed and quality. The findings of this study have implications for the training of translators. Some principles for translation skills training may need to be redefined. In addition to language training, verbal WM training for both TLs and SLs should be introduced, which could enhance students’ translation skills.

However, in this research, only L2→L1 translation has been investigated. As a result, this initial model applies only to L2→L1 translation, not to L1→L2 translation. More research is therefore needed into the effect of verbal WM on translation from L1s into L2s in order to build a complete model for describing the ways in which bilingual WM affects translation.

**References**


model relating memory and lexical retrieval to simultaneous interpreting. *Bilingualism: Language and Cognition, 6*(3), 201–211. https://doi.org/10.1017/S1366728903001135


btl.45.06pac


Appendix I: Questionnaire

1) Would you like to participate in this research project?
   ○ Yes, gladly. ○ No, unfortunately not.

2) Your name: _____________________

3) Your age: _____________________

4) You are
   ○ female ○ male

5) At which university or college are you currently studying? _____________________

6) You are a student
   ○ in the first year of the Master’s degree at the Institute for Translation and Interpreting
   ○ in the second year of the Master’s degree at the Institute for Translation and Interpreting
   ○ in the third year of the Master’s degree at the Institute for Translation and Interpreting

7) Your language level in German corresponds to (multiple selection possible)
   ○ CATTI II. (DE-CN translation) ○ CATTI III. (DE-CN translation)
   ○ Test DaF (with 20 points) ○ Test DaF (with 19 points) ○ Test DaF (with 18 points)
   ○ PGH (very good) ○ PGH (good)

8) Are you used to using online tools such as online dictionaries while translating?
   ○ Yes, I prefer to search and research on the Internet.
   ○ No, I prefer to search and research in print editions.

9) Can you do efficient online research to get relevant information or literature on the Internet while translating?
   ○ Yes, I can. ○ No, unfortunately not.

10) Which operating system do you use?
    ○ Windows ○ MacOS ○ Windows and MacOS ○ Others

11) You are
    ○ Left-handed ○ Right-handed

12) Do you have good eyesight (possibly with glasses)?
    ○ Yes, I can see well (possibly with glasses) ○ No

**Appendix II: Physics quiz on electricity and magnetism**

This quiz includes 40 multiple-choice questions. The following two questions are typical questions taken from the quiz:

1) In telecommunications, a low-frequency useful signal is transmitted to a high-frequency oscillation. This process is called _____.
   
   A. modulation  
   B. frequency modulation (FM)  
   C. amplitude modulation (AM)  
   D. tuning  
   E. no statement

2) Which statement about period duration (T) and frequency (f) applies if the direction of an alternating current changes 100 times within one second?
   
   A. T = 0.01 s  
   B. T = 0.02 s  
   C. f = 100 Hz  
   D. f = 200 Hz  
   E. no statement

**Appendix III: SPM IQ test**

This test includes 60 multiple-choice questions. The following two questions are typical questions taken from the quiz:

1) Which section fits into the picture?

![Diagram C II](image)

A. Section 1  
B. Section 2  
C. Section 3  
D. Section 4  
E. Section 5  
F. Section 6  
G. Section 7  
H. Section 8

2) Which section fits into the picture?

![Diagram D3](image)

A. Section 1  
B. Section 2  
C. Section 3  
D. Section 4  
E. Section 5  
F. Section 6  
G. Section 7  
H. Section 8