Bilingualism: Language and Cognition

cambridge.org/bil

Adaptive control in interpreters: Assessing the impact of training and experience on working memory

Soudabeh Nour[1,2,3], Esli Struys[4,5,6,7] and Helene Stengers[8,9]

[1]Brussels Institute for Applied Linguistics, Vrije Universiteit, Belgium; [2]Centre for Neurosciences, Vrije Universiteit Brussel, Belgium; [3]Brussels Centre for Urban Studies, Vrije Universiteit Brussel, Belgium; [4]Brussels Institute for Applied Linguistics, Vrije Universiteit, Belgium; [5]Centre for Neurosciences, Vrije Universiteit Brussel, Belgium; [6]Brussels Centre for Urban Studies, Vrije Universiteit Brussel, Belgium; [7]Centre for Linguistics, Vrije Universiteit Brussel, Belgium; [8]Brussels Institute for Applied Linguistics, Vrije Universiteit, Belgium and [9]Brussels Centre for Urban Studies, Vrije Universiteit Brussel, Belgium

Research Article

Cite this article: Nour S, Struys E, Stengers H (2020). Adaptive control in interpreters: Assessing the impact of training and experience on working memory. *Bilingualism: Language and Cognition* **23**, 772–779. https://doi.org/10.1017/S1366728920000127

Received: 30 June 2018
Revised: 20 January 2020
Accepted: 22 January 2020
First published online: 17 February 2020

Keywords:
interpreting; working memory; translation; aging; short term memory

Address for correspondence:
Soudabeh Nour, E-mail: Soudabeh.nour@vub.be

Abstract

The adaptive control hypothesis predicts adaptation of control mechanisms as a response to intensive language use in bilinguals. The present study aims to investigate this hypothesis in two memory experiments with professional and student interpreters. In experiment 1, we compared a group of interpreting students to translation students using a reading span task to test working memory (WM) and a digit span task to test short-term memory (STM). In experiment 2, we added a group of professional interpreters and compared them with the participants in experiment 1. Training-related improvement was found for WM but not for STM, with no differences between both student groups. Professional interpreters with over 20 years of interpreting experience showed better performance than translation students but not than interpreting students both on WM and STM. The results are discussed in light of the framework of interpreting as a type of extreme bilingualism.

Introduction

Simultaneous interpreting and working memory

Simultaneous interpreting is known to be a highly demanding cognitive task in which language comprehension and language production happen in approximately the same time in two different languages. Working memory (WM) plays an important role in simultaneous interpreting performance (Daro & Fabbro, 1994) and has become the subject of different studies. As defined by Baddeley (1992) WM is a brain system that is composed not only of the temporary storage of information but also a control system necessary to do complex cognitive tasks (e.g., Conway, Kane & Al, 2005; Jarrold & Towse, 2006). Considering the differences between memory components, researchers have used different span tasks to measure these skills such as simple span tasks (e.g., digit span and word span) for passive storage of information (short-term memory, STM) and complex span tasks (e.g., reading and listening span) for controlling this information (WM). These two memory components are used to a large extent in simultaneous interpreters while they listen to one language and interpret it into another language almost simultaneously. This extensive use of WM during simultaneous interpreting may raise questions such as how interpreting may influence WM in simultaneous interpreters, whether simultaneous interpreting practice could enhance WM, or if having a good WM is a prerequisite to be a good simultaneous interpreter?

The realisation that the naturalistic task of language interpretation relies on WM can be related to the WM advantage in interpreters which refers to the finding that groups of interpreters outperform non-interpreters on WM tasks (De Groot & Christoffels, 2006; Signorelli, Haarmann & Obler, 2012; Babcock & Vallesi, 2017). Two recent meta-analyses on memory research with interpreters confirmed that this advantage affects both WM and STM spans and that the size of this advantage is moderated by the level of experience in interpreting (Mellinger & Hanson, 2019; Wen & Dong, 2019). The presence of an interpreter advantage in WM can be explained in light of the adaptive control hypothesis (Green & Abutalebi, 2013), which predicts adaptation of control processes according to the interactional context in which a bilingual is immersed. This interactional context can be either a single-language, a default dual-language or a dense dual-language (code-switching) context. In ordinary bilingual language use, dual-language contexts are composed of a sequential alternation between the two (or more) languages with no or very little overlap between the languages. As simultaneous interpreters are frequently exposed to a context with continuous overlap between the two languages, the control demands of simultaneous interpreting are even higher than

© The Author(s) 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

CAMBRIDGE UNIVERSITY PRESS

those of dual-language contexts in ordinary bilingual language use which is why simultaneous interpreting has been labelled as an extreme form of bilingualism that requires 'extreme language control' (e.g., Hervais-Adelman, Moser-Mercer, Michel & Golestani, 2014). Within the framework of the adaptive control hypothesis, the finding of an interpreter advantage on WM may thus be seen as an example as to how one specific context (i.e., simultaneous interpreting) leads to 'extreme' adaptation of control abilities in those individuals with yearlong exposure to this context (Christoffels, de Groot & Kroll, 2006; Henrard & Van Daele, 2017; Yudes, Macizo & Bajo, 2011). However, these advantages for interpreters have not been found consistently (Babcock, Capizzi, Arbula & Vallesi, 2017; Stavrakaki, Megari, Kosmidis, Apostolidou & Takou, 2012; Tzou, Eslami, Chen & Vaid, 2012). In the following paragraphs, we will discuss three important issues in the current literature on the interpreting advantage in WM that could explain these contradictory findings.

A first important issue in the study of the interpreter advantage in WM concerns the characteristics of the interpreter group showing the advantage. Better performance for interpreters on WM has both been found in interpreter trainees and interpreting professionals (Dong & Liu, 2016; Henrard & Van Daele, 2017; Signorelli et al., 2012), which leads to the question of how much interpreting experience would be needed to show the advantage. In this respect, the age factor is an important interfering variable that must be taken into account in any research of the effects of interpreting experience and training. Evidently, experienced interpreters tend to be older than interpreter trainees who are typically in their early twenties which corresponds to the optimal age of cognitive functioning (Chmiel, 2018; Köpke & Nespoulous, 2006). Importantly, normal aging is associated with a decline in WM performance (Fabiani, 2012). It is revealed that short-term intensive language training has a positive effect on cognitive components such as attention network in a healthy older age population (Bak, Long, Vega-Mendoza & Sorace, 2016). Additionally, research in bilingualism and aging has revealed that cognitive decline is delayed in bilinguals compared to monolinguals with Alzheimer's disease (Bialystok, Abutalebi, Bak, Burke & Kroll, 2016; Bialystok, Craik & Freedman, 2010). In the case of interpreters, the professional interpreting experience can be seen as an adaptive control process in a highly demanding dual language context. According to the literature, this adaptation connected to interpreting leads to a cognitive advantage for professional interpreters compared to non-interpreting groups that is even more pronounced at older age than in early adulthood (Chmiel, 2018; Henrard & Van Daele, 2017). These findings seem to suggest that aging-related decline in WM may be compensated for by the protective influence of adaptive control in professional interpreters.

A second important issue in the study of the relationship between interpreting and memory is the question of exactly which memory components are affected by the interpreting experience. Previous literature showed that interpreting affects not only WM as tested by a reading span test (Signorelli et al., 2012; Yudes et al., 2011) but also STM as tested by a digit span test (Tzou, Eslami, Chen & Vaid, 2016). From an adaptive control perspective, it is logical to assume that specifically WM (but not STM) will show adaptation as a result of continuous and simultaneous exposure to both languages because only the former component taps into control processes. In light of this, it is remarkable that studies showing an interpreter advantage on STM have only used tasks with linguistic stimuli (such as letters, and words) (Babcock et al., 2017), while studies showing a similar advantage on WM have used a variety of tasks with different stimuli and modalities such as the n-back task using a visuospatial version of the 2-back with blue squares as stimuli, and the operation and reading span tests (Dong & Liu, 2016; Morales, Yudes, Gómez-Ariza & Bajo, 2015; Signorelli et al., 2012).

A final issue in the study of WM and interpreting is the extent to which observed effects can be generalised across languages. Many studies have used WM tasks in L1 and L2 to test interpreters (e.g., Chmiel, 2018; Tzou et al., 2016). Most of these studies have reported larger WM and STM spans in L1 than L2 in groups of interpreter including both students and professionals (Cai, Dong, Zhao & Lin, 2015; Chmiel, 2018; Tzou et al., 2016) except for one study (Christoffels et al., 2006) which reported no significant differences for WM complex spans in L1 vs. L2 in professional interpreters. These divergent results can be explained by the specificities of the interpreting profession, most notably in relationship to language directionality. As most interpreters tend to interpret more frequently from their second language into their mother tongue or most proficient language, it can be expected that they are more experienced in controlling WM representations from the input language, i.e., L2. This could have an effect on the adaptation of control skills in interpreting professionals, possibly leading to more extensive L2 control skills.

The present study

The goal of the present study was to investigate in two experiments to what extent the adaptive control hypothesis (Green & Abutalebi, 2013) can be applied to WM performance in interpreters (students and professionals) and translation students (control group). We tested the effects of interpreting training and experience by investigating two groups of trainees with no previous experience in interpreting (or translation) and a group of professionals with years of relevant practice. Based on the assumption that interpreting is an extreme form of bilingual language control (Hervais-Adelman et al., 2014), we expected a specific effect of this special language activity on WM but not on STM.

In the first experiment, we assessed the effect of interpreter training on WM and STM using a longitudinal design. The general approach we took was the same as that in the longitudinal studies by Babcock et al. (2017) and Dong and Liu (2016) which compared the performance of interpreter students with a control group of students on two different time points. Pre- and post-training comparison between the groups allowed us to establish if potential group differences are specifically related to the training that the interpreters receive or if they could be ascribed to pre-existing individual differences related to self-selection of interpreter students (Babcock et al., 2017). In contrast to the two previous studies using a similar research design (Babcock et al., 2017; Dong & Liu, 2016), we selected WM tasks that allowed us to test the generalisability of adaptive control in interpreters over languages and language modality. The reading span test was chosen for because it involves the written language modality, which is not directly involved in interpreting as an oral activity. Although some written input may be available in the interpreting process, the main language input is still in the audio mode. Additionally, the standard reading span test is available in different languages which makes it possible to assess comparable WM scores in L1 and L2 especially in our group of

https://doi.org/10.1017/S1366728920000127 Published online by Cambridge University Press

professional interpreters. This test was taken in two languages to test the hypothesis of language-independence, or the question whether adaptive control applies to both languages of the interpreter. We hypothesised that interpreting training may affect WM differently from translation training taking into account the language of the task (L1 vs. L2). Firstly, WM has a time limitation in the sense of information storage and processing; likewise, interpreting involves immediacy that is not present during translation. A recent longitudinal study by Dong, Liu and Cai (2018) revealed that consecutive interpreting students show more progress than general L2 students in WM using an n-back task. We hypothesised that the immediacy factor may lead to better WM performance in interpreting groups. Especially we predicted a better performance (or more progress) for interpreter students in the reading span test in L2 given that the training that these students received mainly involves L2-to-L1 interpreting, so with L2 as the input language.

In the second experiment, we assessed the effect of interpreting experience on WM and STM using a cross-sectional design. We added the professional interpreter group in line with the study design of Chmiel (2018) and as suggested by Babcock et al. (2017) to identify if any group differences could be found when we compared both young students groups (from the first experiment) with older professional interpreters with a big age gap between them, and if there is any decline in professional interpreters' WM performance due to the age factor that we can see in normal populations (Fabiani, 2012). The study by Chmiel (2018) compared professional interpreters with interpreting students at pre- and post-training and with non-interpreting bilingual students only at training onset. The novelty of the present study lay in the comparison of professional interpreters' WM performance not only with pre- but also with post-training performance of both interpreting students and control translation students from the first experiment. In this way, we could test if the professional advantage over young non-interpreting group as reported by Chmiel (2018) persisted after training. This addition was crucial to untangle potential effects of adaptive control related to training and experience. In line with the results from the recent meta-analyses by Mellinger and Hanson (2019) and Wen and Dong (2019), we predicted that interpreting experience may positively affect memory components across the life span, and serves as a protective buffer against the detrimental effects of aging on cognitive performance.

Method

Experiment 1: Longitudinal study

In this experiment two groups of interpreting students and translation students were tested for their WM function using two different tasks. A reading span test was administered in their first and second language (L1 and L2) to test WM, and a digit span test was administered to test the participants' STM capacity before and after their one-year Master's programme.

Participants

A total of thirty-eight students from the Dutch-medium Vrije Universiteit Brussel in Belgium (29 females) participated in this longitudinal experiment. Based on their one-year Master's programme, the population was further subdivided into two groups: the translation students and the interpreting students. Both groups received their Bachelor's degree in applied linguistics before entering the Master's programme. The first group consisted of 17 students (15 females) with a mean age of 22.2 years ($SD = 1.8$), pursuing a Master's degree in interpreting. The second group consisted of 21 students (14 females) with a mean age of 23.1 years ($SD = 2.9$), earning a Master's degree in translation. The students received either course credit or reimbursement for their participation in the tests.

The Master's programme in interpreting consists of theoretical courses into interpreting studies, research methodology and intercultural communication (a total of 4 hours a week), and interpreting training in class focusing on memory exercises, consecutive, simultaneous and sight interpreting from and into Dutch (as the A-language) for a total of 4 hours a week per target language. The interpreting classroom of the Vrije Universiteit Brussel contains 14 booths and the students spend most of the time in these cabins, especially in the second semester of the programme. Students also go on an intensive interpreting internship in a multilingual company or public institution (e.g., the Belgian parliament) where they are immersed in a real-life interpreting context. In light of one of the tasks that was used in our test battery, it is important to note that interpreting trainees are trained to remember digits easily and convert them accurately and quickly into another language. This is particularly important because of the incongruence between languages that put the unit before the ten such as Dutch and German (e.g., 21 in Dutch is éénentwintig, lit. translated: "one and twenty"), and languages that do the inverse such as English and French. Interpreter trainees are taught specific techniques and strategies to smooth the translation of number words and to reverse the digits in case of incongruence between the involved languages. The Master's programme in translation is built up analogously to the Master's programme in interpreting, meaning that it offers a mixture of theoretical courses into translation studies, research methodology and intercultural communications (a total of 4 hours a week) and translation training in class composed of workshops on translation strategies and techniques and guided assignments for a total of 2 hours a week per target language. In addition, students are required to complete assignments at home and they receive extensive and individual feedback on these assignments from the translation teachers. Just like in the Master's programme in interpreting, translation trainees go on an intensive internship to learn more about the translation profession. Importantly, both student groups have to choose at least two languages (one B-language) during their Master's programme.

In order to collect information about L1 and L2 proficiency, the two student groups completed an adapted version of the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld & Kaushanskaya, 2007) in Dutch including questions about the number of languages they spoke, their onset ages of language acquisition for L1 and L2, self-reported interpreting or translation proficiency on a 10-point scale, years of interpreting and translation experience, and exposure to the languages in the twelve months preceding the time of investigation (in percentages). Even though all participants were enrolled in a Dutch-speaking Master's programme, some participants self-reported to have another L1. Out of the 38 participating students, 33 had Dutch as their L1 and 5 French. Twenty-three students reported French as their second language, 13 had English as their L2 and 2 Dutch. Details of the language background characteristics of the participants are presented in Table 1.

Inferential statistics were conducted to assess to what extent the three groups differed from each other on these background

https://doi.org/10.1017/S1366728920000127 Published online by Cambridge University Press

Table 1. Language background characteristics of participants

	Translation student Means (SD)	Interpreting student Means (SD)	Professional interpreter Means (SD)
Age	23.11 (2.95)	22.28 (1.8)	52.73 (6.85)
AoA L1	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
AoA L2	6.78 (4.45)	5.44 (4.21)	8.03 (4.1)
Recent exposure L1	50.26 (18.83)	47.12 (14.51)	43.85 (17.36)
Recent exposure L2	22.53 (12.73)	16.71 (10.33)	19.21 (17.19)
TRA/INT into L1 (self-rated proficiency)	7.67 (0.71)	6.41 (1.41)	9.07 (0.45)
TRA/INT experience (pre-test)	0.00 (0.00)	0.00 (0.00)	24.60 (12.15)

L1: first language, L2: second language, AoA: age of acquisition expressed in years, recent exposure in L1 and L2 in the twelve months preceding the time of investigation expressed in percentages, TRA: translation, INT: interpreting, self-rated proficiency expressed on a 10-point scale; TRA/INT experience (pre-test): translation or interpreting experience before first test session expressed in years.

characteristics. Independent samples' T-tests did not reveal any differences between the groups on any of the self-reported measures (all $p > .05$). No tests were conducted on the onset age of L1 acquisition because it was (nearly) the same for all participants.

Material and tasks

In the digit span test, participants were presented with random strings of numbers. The length of the number string increased on each repetition with a set of two digits at the beginning and a set of 9 numbers at the end. The participants had to remember the order of the numbers they heard in each section. The stimuli for the digit span task were digits from 1 to 9 and they were presented in the auditory mode. Each trial started with a fixation ('+') on the centre of a screen followed by the audio presentation of string of digits. The duration of each digit was 1000 milliseconds, with a 500 milliseconds interval between two successive digits. After the presentation of the last digit in each set, participants were asked to recall the digits in the order they heard by pressing the upper number keys on the keyboard. If the participants failed to recall two successive trials of a given set correctly the digit span test was stopped automatically. The participant's digit span was the longest string of numbers which was correctly recalled with a maximum score of nine. E-Prime software 2.0 (Schneider, Eschman & Zuccolotto, 2012) was used for test design.

For the reading span test (Daneman & Carpenter, 1980), we used the shortened version which was developed by Van Den Noort, Bosch, Haverkort and Hugdahl (2008) (60 rather than 100 sentences), both in the first and second language of the participants. The advantage of this version of the reading span test is that the test variants were designed in different languages, and in each language the sentences are matched for length and word frequency within and across languages. The reading span test versions which are used in this study are Dutch, French, German and English based on the participants' language background (L1 and L2) leading to more reliable and comparable data across different languages. During the task participants were presented with three blocks, and each block contained 20 sentences (60 sentences in total). In each block, a series of sentences contained two, three, four, five, or six sentences. The participants were asked to read out loud the sentences presented on the monitor while remembering each sentence-final word. At the start of the tests participants received an oral instruction and were presented with a written explanation on the screen during the first two sentences which were trial sentences. At the end of each series of sentences, the word "RECALL" (in the language of the test) appeared on the screen asking participants to recall as many sentence-final words as they could remember considering that the order of the words they recalled was not important. In line with the recommendations provided by Van den Noort et al. (2008), free recall was allowed to avoid primacy and recency effects. Each correctly recalled word was awarded by one score; the maximum score was 60 or the total number of sentences in the test (for more details about the scoring method, please see Friedman & Miyake, 2005). To minimise the repetition task effect, different versions of the reading span test both in L1 and L2 were administered in pre-training and post-training test sessions of this longitudinal study. E-Prime software 2.0 (Schneider et al., 2012) was used for test design.

Procedure

All participants were tested in the behavioural lab at the Department of Psychology and Educational Sciences of the Vrije Universiteit Brussel (VUB). The lab has 14 separate soundproof cabins; each cabin contains a button box, a monitor, a microphone, a pair of headphones, a keyboard and a recorder. The total time needed for finishing the tasks was between 20 to 25 minutes according to the participants' performance rate. The participants received test instructions both orally (by instructor) and in written form (through monitor) before starting each test. These students were tested at the start of their Master's programme and at the end of the programme with a nine months interval between both measurements. The university's guidelines regarding ethical research and scientific integrity were strictly followed. All participants gave informed consent for participation to this experiment.

Experiment 2: Cross-sectional study

In this experiment, professional interpreters were tested on their memory performance to investigate the effect of interpreting experience. The goal of this experiment was to better understand to what extent the professional career of interpreting may influence WM and STM in this expert population, and to what extent it could protect them against cognitive decline as a result of normal aging processes. To this aim we compared memory performance of professional interpreters with memory performance of both interpreting and translation students from the first experiment at pre- and post-training separately.

Participants

This group was composed of 21 professional conference interpreters (11 females) with a mean age of 52.73 years ($SD = 6.85$), recruited voluntarily from the Directorate-General for Interpretation (DG Interpretation) of the European Commission in Brussels by answering the open call that was posted on the internal website of the DG Interpretation. They were tested only on one time point.

https://doi.org/10.1017/S1366728920000127 Published online by Cambridge University Press

Professional interpreters completed an adapted version of the Language Experience and Proficiency Questionnaire, LEAP-Q (Marian, Blumenfeld & Kaushanskaya, 2007) in English, including questions about the number of languages they spoke, their onset ages of language acquisition for L1 and L2, self-reported interpreting proficiency on a 10-point scale, interpreting experience in years, and exposure to their languages in the twelve months preceding the time of investigation (in percentages). Details of the participants' background information are presented in Table 1.

Inferential statistics were conducted to assess to what extent the three groups differed from each other on these background characteristics. Analyses of variance (ANOVA) did not reveal any differences between the three groups on the onset age of L2 acquisition, and on recent exposure to L1 and L2 (all $p > .05$). As could be expected, between-group differences were found on age, $F(2, 57) = 266.35$, $p < .001$, and on self-reported proficiency in translation or interpretation, $F(2, 57) = 139.35$, $p < .001$. In both cases, post-hoc Bonferroni-corrected tests revealed that the professional interpreters were significantly different from the two other groups ($p < .001$), who did not differ from each other. The professional interpreters had a higher age and a higher self-reported proficiency in translation or interpretation (for the mean scores, see Table 1). No tests were conducted on the onset age of L1 acquisition because it was (nearly) the same for all participants, and on professional experience, because all students reported no (or very little) experience.

Material and tasks

The materials used in this experiment including digit span and reading span test (L1 and L2) were the same as in Experiment One. All tests were designed with E-Prime software 2.0 (Schneider et al., 2012).

Procedure

The same procedures were followed as in Experiment One. All participants were tested in the behavioural lab at the Department of Psychology and Educational Sciences at Vrije Universiteit Brussel (VUB). The university's guidelines regarding ethical research and scientific integrity were strictly followed. All participants gave informed consent for participation to this experiment.

Results

Experiment 1: Longitudinal study

Digit span

The descriptive statistics of the digit span scores can be found in Table 2. The digit span scores of the two student groups did not differ significantly from each other before training, $t(36) = 1.05$, $p > .05$. A two-way ANOVA with Time as a within-subject variable and Group as a between-subject variable was performed on the digit span scores. Both independent variables had two levels: for Time these levels were pre- and post-training, for Group these were interpreting and translation trainees. We found no main effect of Time, no main effect of Group, and no interaction effect between Time and Group, all $p > .05$.

Reading span

The descriptive statistics of the reading span scores can be found in Table 2. The reading span scores of the two student groups did not differ significantly from each other before training, neither for L1, $t(36) = 0.97$, $p > .05$; nor for L2, $t(36) = 0.52$, $p > .05$. A three-way ANOVA with Time and Language as within-subject variables, and Group as a between-subject variable was performed on the reading span scores. All three independent variables had two levels: for Time, these levels were pre- and post-training, for Language these were L1 (including Dutch = 33, French = 5) and L2 (including Dutch = 2, French = 23, and English = 13), and for Group these were interpreting and translation trainees. The results showed a main effect of Time, $F(1,25) = 39.67$, $p < .001$, $\eta p2 = .61$, with higher scores on the reading span test in post training for interpreters (Pre: $M = 36.1$, $SD = 1.75$; Post: $M = 41.63$, $SD = 1.8$) as well as for translators (Pre: $M = 32.91$, $SD = 1.95$, Post: $M = 37.12$, $SD = 2.00$). Additionally, a main effect of Language, $F(1,25) = 38.96$, $p < .001$, $\eta p2 = .61$, was found for both groups of students with higher scores for L1 than L2 in interpreters (L1: $M = 40.80$, $SD = 1.71$; L2: $M = 36.93$, $SD = 1.76$) and translation students (L1: $M = 36.83$, $SD = 1.91$; L2: $M = 33.20$, $SD = 1.97$), but no main effect of Group, $F(1,25) = 2.29$, $p > .05$, $\eta p2 = .08$, no two-way interaction effects between Time and Language, Time and Group, and Language and Group, and also no three-way interaction effect between Time, Language and Group were found (all $p > .05$).

Experiment 2: Cross-sectional study

Digit span

The descriptive statistics of the digit span scores can be found in Table 2. A one-way ANOVA with Group as a between-subject variable was performed on the digit span scores to compare the scores of the professional interpreters to those of the interpreting and translation trainees before their training. The group variable thus had three levels. We found no main effect of Group, $p > .05$. Another one-way ANOVA with Group as a between-subject variable was performed on the digit span scores to compare the same scores of the professional interpreters to the scores of the interpreting and translation trainees after their training. This second analysis revealed a main effect of Group, $F(2,45) = 3,30$, $p < .05$, $\eta p2 = .13$. Post hoc comparisons using the Tukey HSD test revealed that only the difference between professional interpreters and translation students ($M = .82$, $SD = .29$, $p < .05$) was significant due to the better performance of professional interpreters ($M = 7.57$, $SD = 0.75$) compared to translation students ($M = 6.75$, $SD = 8.67$). All other pairwise comparisons did not reveal any further significant difference, all $p > .05$.

Reading span

The descriptive statistics of the reading span scores for professional interpreters can be found in Table 2. A two-way ANOVA with Language as a within-subject variable, and Group as a between-subject variable was performed on the reading span scores to compare the scores of the professional interpreters to those of the interpreting and translation trainees before their training. Languages of the reading span test for professional interpreters were L1 (including: Dutch = 2, German = 6, French = 5 and English = 8) and L2 (including: Dutch = 2, German = 2, French = 7 and English = 10). We found a main effect of Group, $F(2,56) = 2.25$ $p < .05$, $\eta p2 = .16$, and a main effect of Language, $F(2,56) = 55.61$, $p < .001$, $\eta p2 = .50$, with higher scores on L1 ($M = 39.51$; $SD = 1.20$) than on L2 ($M = 35.38$; $SD = 1.03$), but no two-way interaction effect between Language and Group was

https://doi.org/10.1017/S1366728920000127 Published online by Cambridge University Press

Table 2. Digit span and reading span scores for both student groups (pre- and post-training) and for professional interpreters

	Translation student Means (SD)	Interpreting student Means (SD)	Professional interpreter Means (SD)
Digit span (pre-training)	7.00 (1.38)	7.41 (0.94)	n/a
Digit span (post-training)	6.75 (0.86)	7.20 (0.86)	7.57 (0.75)
Reading span L1 (pre-training)	37.38 (7.44)	39.88 (8.44)	n/a
Reading span L1 (post-training)	39.75 (5.99)	43.80 (6.96)	44.57 (6.78)
Reading span L2 (pre-training)	33.43 (7.05)	34.65 (7.39)	n/a
Reading span L2 (post-training)	33.92 (7.64)	37.80 (7.58)	39.81 (7.27)

Mean scores are reported with standard deviations between brackets. n/a = not applicable. The digit span score refers to the longest string of numbers which was correctly recalled with a maximum score of nine. The reading span scores refer to the number of correctly recalled words with a maximum score of 60.

found, $p > .05$. Post hoc comparisons using the Tukey HSD test revealed that the mean difference between professional interpreters and translation students ($M = 6.79$; $SD = 2.15$, $p < .05$) was significant due to the better performance of professional interpreters compared to translation students both in L1 and L2 at pre-training. Another two-way ANOVA with the same independent variables was performed on the reading span scores to compare the same scores of the professional interpreters to the scores of the interpreting and translation trainees at post-training. This second analysis revealed a main effect of Language, $F(2,45) = 61.75$, $p < .001$, $\eta p2 = .58$, with higher scores for L1 ($M = 40.70$; $SD = 0.99$) than for L2 ($M = 37.17$; $SD = 1.10$), but no main effect of Group, and no interaction effect between language and group, all $p > .05$.

Discussion

The present study intended to test the adaptive control hypothesis (Green & Abutalebi, 2013) in interpreter trainees and professional interpreters with the aim to untangle potential effects of interpreter training and experience on WM and STM. Our overarching study design included both a longitudinal research with students of translation as a control group to investigate training effects and a cross-sectional research with professional interpreters to study the impact of experience.

This study did not detect any differences between interpreting students and translation students at pre-training. This result is in line with other longitudinal studies which reported no significant pre-existing advantage for interpreting students compared to translation students (Babcock et al., 2017; Dong & Liu, 2016). Furthermore, our results revealed a main effect of time selectively on WM as tested by the reading span test but not on STM as tested by the digit span test. The selective effect of interpreter and translator training on WM and not on STM seem to confirm the adaptive control hypothesis, even though it must be admitted that these different results for WM and STM tasks could be due to the nature of the materials which are more meaningful for the reading span task and may therefore be more amenable to practice effects. However, we expected to find a larger effect in the interpreters than in the translators, but this expectation was not borne out by the data: we did not find any difference at post-training or in the degree of improvement both in WM L1 and WM L2 between the interpreting and translation trainees. Our findings rather suggest that both types of training rely on similar degrees of control but, to exclude the possibility that the observed improvement is a result of the simple fact that both groups were one year older at the post test, this can only be confirmed by future studies that compare interpreting or translating trainees to a third (control) group that underwent a different type of training. Possibly, the contrast between simultaneous interpreting and translation regarding control requirements is too little to see any significant difference in WM improvement at least after a one-year Master's programme.

Our results on WM are in line with previous longitudinal studies that had a similar research design by Dong and Liu (2016) and Babcock et al. (2017), who found improvement for interpreting and translation students on WM, using the operation span and n-back test. The results of the present study are different from those reported by Macnamara and Conway (2014), who found no improvement in interpreting and translation students on WM, also using an operation task. We believe that these contradictory results may come back to specific characteristics of the operation span task. The validity of the operation span task was put into question in a recent study which reported that this task is not as sensitive in distinguishing between normal and above average populations on WM tasks compared to other tasks such as the symmetry span and rotation span tasks (Draheim, Harrison, Embretson & Engle, 2018). It should be noted that, while Dong and Liu (2016) found significant improvement for the interpreters and marginally significant improvement for the translators on WM, the current study found similar improvement for both groups on the same construct. One possible reason for this difference could be that the study by Dong and Liu (2016) compared the performance of translators and interpreters to a control group, while the current study did not include such a group. A second reason for this difference may be found in the usage of a reading span task in the present study to assess WM. As translators tend to work with written texts as input, it could be that their specific training in controlling two language systems while reading texts in the source language could have boosted their performance on the reading span test in the present study.

Our results on STM are in line with the cross-sectional study by Tzou et al. (2012) who found no difference on digit span scores between first-year and second-year students. However, two longitudinal studies by Babcock et al. (2017) and Macnamara and Conway (2014) showed a different result; they reported training-related improvement in interpreting students using a letter span task. One reason for this replication failure may come back to the length of training which was one academic year in the present study, but two years in the study by Babcock et al. (2017). This may lead to the idea that STM needs more training or experience to show its advantage in interpreting groups at least in behavioral level.

Regarding the impact of experience, we found that professional interpreters with substantial experience were significantly better on WM than students of translation in both L1 and L2 at pre-training. Interestingly, the same group comparison with students of interpreting and professional interpreters turned out not to be significant, even though group effects between translator and interpreter students were not detected in the longitudinal study.

https://doi.org/10.1017/S1366728920000127 Published online by Cambridge University Press

The results of the professional interpreters are remarkable, taking into account the time gap of more than 30 years between the group of professional interpreters and the student groups. The results suggest that the accumulation of interpreting experience can keep WM components in their optimal form and may compensate for the cognitive decline related to normal aging (Henrard & Van Daele, 2017). The results of the current study are in line with those by Chmiel (2018) who reported better performance for professional interpreters compared to a group of bilingual students but not over interpreting students, and those by Henrard and Van Daele (2017), who reported that group differences between professional interpreters and other control groups (including translators) became larger with increasing age. They suggested that accumulated experience in simultaneous interpretation may be the reason for these differences. In light of the adaptive control hypothesis, these studies and ours provide some evidence that WM undergoes long-term adaptation to meet the demands of daily interpreting as a form of extreme language control at a professional level, even though it must be stressed that we could not find significant differences between interpreting professionals and trainees, and that a definitive conclusion on this issue can only be drawn from a study where interpreting professionals are compared to an age-matched control group of translators with the same amount of experience.

Our study revealed higher performance of professional interpreters in STM compared to translation students at post-training. Analogous to the pattern of results on WM, the group effect was not found when comparing the professional group to the interpreting students, again pointing to subtle between-group differences in the longitudinal study that did not reach the level of statistical significance. This pattern of results confirms that professional experience is needed to see any effects of interpreting on STM. Moreover, this impact of experience of WM capacity seems to generalise over domains as the digit span task used in the current study was not composed of linguistic but of numerical stimuli.

For the first time, we tested interpreters on WM tasks longitudinally in different languages (L1 and L2) to find out more about the interaction of language specificity and WM and how language specificity may affect control adaptation in interpreters. No different effects were found for L1 and L2, which may be remarkable given the fact that the training these interpreters received is predominantly oriented towards the interpretation of oral texts from a second language into the native (or most proficient) language. These results are in line with the previous longitudinal study by Chmiel (2018) who reported an improvement in reading span test scores in a second language for interpreting students after training. These results suggest that any training effect related to WM may be generalisable to all working languages involved in interpreting and not only to the language that is used as the input language.

Even though the present study introduces new insights into the application of the adaptive control hypothesis to the field of interpreting, there are a few limitations which need to be addressed. Firstly, the scope of the present study is limited due to the small number of participants. Secondly, to better understand the long-term effect of dual language experience on WM it could have been better to have added a control group for professional interpreters. Further studies with a focus on the contrast between active and non-active interpreters in later ages could be more informative to find out if the WM advantage of interpreters is caused by individual differences or rather the result of an accumulation of interpreting experience; for example, by comparing professionally active interpreters to their peer interpreters from the same age who recently quit the job and who had become professionally inactive as a consequence. If any differences are found between the two groups, one can be more confident that it is the high WM activity which is required during simultaneous interpreting that is the driving force behind the WM enhancement and not pre-existing individual differences. Thirdly, as mentioned in the introduction we should distinguish between behavioural studies and neuroimaging studies. Van de Putte et al. (Van de Putte, De Baene, García-Pentón, Woumans, Dijkgraaf & Duyck, 2018) reported no behavioral differences between interpreting and translation students after one year of university training; however, changes in brain plasticity and functional networks were detected. Future studies in interpreting should include different techniques while using different kinds of WM tasks in the same study to look at the role of training and experience in interpreting groups.

Conclusion

The results of the present study shed some light on the adaptive control hypothesis in the context of professional interpreters and interpreting students. Training-related improvement was selectively seen on WM but not on STM, but there were no differences between interpreting and translation trainees in this respect. Moreover, interpreting professionals with yearlong expertise in the profession only showed enhanced WM performance compared to much younger students of translation, but not to students of interpreting. This study provides also a few new insights. First, the progress on WM was seen as well in interpreting as in translation students, so there does not seem to be a unique effect of interpreting versus translation training on WM performance. Second, the better performance of interpreting professionals compared to much younger students of translation is remarkable but conclusive evidence about the impact of long-term interpreting experience can only be collected from studies where interpreting professionals are compared to an age-matched control group of professional translators. The present study can serve as a starting point for a further exploration of the adaptive control hypothesis in the interpreting domain: for instance, by comparing WM performance of professionally active interpreters to that of their inactive peers of the same age using both behavioural and neuroimaging techniques.

Acknowledgements. We are grateful to the Department of Linguistics and Literary Studies, section Applied Linguistics, of the Vrije Universiteit Brussel for their generous support in funding the present research (reference code = ACA1).

References

Babcock, L, Capizzi, M, Arbula, S, & Vallesi, A (2017). Short-Term Memory Improvement After Simultaneous Interpretation Training. *Journal of Cognitive Enhancement*, **1**, 254–267. https://doi.org/10.1007/s41465-017-0011-x

Babcock, L, & Vallesi, A (2017). Are simultaneous interpreters expert bilinguals, unique bilinguals, or both? *Bilingualism: Language and Cognition*, **20**, 403–417. https://doi.org/10.1017/S1366728915000735

Baddeley, A (1992). Working memory. *Science*, **255**, 556–559. https://doi.org/10.1126/science.1736359

https://doi.org/10.1017/S1366728920000127 Published online by Cambridge University Press

Bak, TH, Long, MR, Vega-Mendoza, M, & Sorace, A (2016). Novelty, challenge, and practice: The impact of intensive language learning on attentional functions. *PLoS ONE*, **11**. https://doi.org/10.1371/journal.pone.0153485

Bialystok, E, Abutalebi, J, Bak, TH, Burke, DM, & Kroll, JF (2016). Aging in two languages: Implications for public health. *Ageing Research Reviews*, **27**, 56–60. https://doi.org/10.1016/j.arr.2016.03.003

Bialystok, E, Craik, FIM, & Freedman, M (2010). Delaying the onset of Alzheimer disease. *Neurology*, **75**, 1726–1729. https://doi.org/10.1212/WNL.0b013e3181fc2a1c

Cai, R, Dong, Y, Zhao, N, & Lin, J (2015). Factors contributing to individual differences in the development of consecutive interpreting competence for beginner student interpreters. *The Interpreter and Translator Trainer*, **9**, 104–120. https://doi.org/10.1080/1750399X.2015.1016279

Chmiel, A (2018). In search of the working memory advantage in conference interpreting - Training, experience and task effects. *International Journal of Bilingualism*, **22**, 371–384. https://doi.org/10.1177/1367006916681082

Christoffels, IK, de Groot, AMB, & Kroll, JF (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, **54**, 324–345. https://doi.org/10.1016/j.jml.2005.12.004

Conway, ARA, Kane, MJ, & Al, CET (2005). Working memory span tasks : A methodological review and user ' s guide. *Psychonomic Bulletin & Review*, **12**, 769–786. https://doi.org/10.3758/BF03196772

Daneman, M, & Carpenter, PA (1980). Individual differences in working memory during reading. *Journal Of Verbal Learning And Verbal Behavior*, **19**, 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Daro, V, & Fabbro, F (1994). Verbal memory during simultaneous interpretation - Effects of phonological interference. *Applied Linguistics*, **15**, 365–381. https://doi.org/10.1093/applin/15.4.365

De Groot, AMB, & Christoffels, IK (2006). Language control in bilinguals: Monolingual tasks and simultaneous interpreting. *Bilingualism*, **9**, 189. https://doi.org/10.1017/S1366728906002537

Dong, Y, & Liu, Y (2016). Classes in translating and interpreting produce differential gains in switching and updating. *Frontiers in Psychology*, **7**. https://doi.org/10.3389/fpsyg.2016.01297

Dong, Y, Liu, Y, & Cai, R (2018). How does consecutive interpreting training influence working memory: A longitudinal study of potential links between the two. *Frontiers in Psychology*, **9**, 1–12. https://doi.org/10.3389/fpsyg.2018.00875

Draheim, C, Harrison, T, Embretson, S, Engle, R (2018). What Item Response Theory Can Tell Us About the Complex Span Tasks. *Psychological Assessment*, **30**, 116–129. https://doi.org/10.1037/pas0000444

Fabiani, M (2012). It was the best of times, it was the worst of times: A psychophysiologist's view of cognitive aging. *Psychophysiology*, **49**, 283–304. https://doi.org/10.1111/j.1469-8986.2011.01331.x

Friedman, NP, & Miyake, A (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods*, **37**, 581–590. https://doi.org/10.3758/BF03192728

Green, DW, & Abutalebi, J (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**, 1–16. https://doi.org/10.1080/20445911.2013.796377

Henrard, S, & Van Daele, A (2017). Different Bilingual Experiences Might Modulate Executive Tasks Advantages: Comparative Analysis between Monolinguals, Translators, and Interpreters. *Frontiers in Psychology*, **8**, 1–11. https://doi.org/10.3389/fpsyg.2017.01870

Hervais-Adelman, A, Moser-Mercer, B, Michel, CM, & Golestani, N (2014). fMRI of Simultaneous Interpretation Reveals the Neural Basis of Extreme Language Control. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhu158

Jarrold, C, & Towse, JN (2006). Individual differences in working memory. *Neuroscience*, **139**, 39–50. https://doi.org/10.1016/j.neuroscience.2005.07.002

Köpke, B, & Nespoulous, J (2006). Working memory performance in expert and novice interpreters. *Most*, **1**.

Macnamara, BN, & Conway, ARA (2014). Novel evidence in support of the bilingual advantage: Influences of task demands and experience on cognitive control and working memory. *Psychonomic Bulletin & Review*, **21**, 520–525. https://doi.org/10.3758/s13423-013-0524-y

Marian, V, Blumenfeld, H, & Kaushanskaya, M (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing research*, **50**, 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mellinger, C, & Hanson, T (2019). Meta-analyses of simultaneous interpreting and working memory. *Interpreting*, **21**, 165–195. https://doi.org/10.1075/intp.00026.mel

Morales, J, Yudes, C, Gómez-Ariza, CJ, & Bajo, MT (2015). Bilingualism modulates dual mechanisms of cognitive control: Evidence from ERPs. *Neuropsychologia*, **66**, 157–169. https://doi.org/10.1016/j.neuropsychologia.2014.11.014

Schneider, W, Eschman, A, & Zuccolotto, A (2012). E-Prime 2.0 Reference Guide Manual. *Psychology Software Tools, Inc.*

Signorelli, TM, Haarmann, HJ, & Obler, LK (2012). Working memory in simultaneous interpreters: Effects of task and age. *International Journal of Bilingualism*, **16**, 198–212. https://doi.org/10.1177/1367006911403200

Stavrakaki, S, Megari, K, Kosmidis, MH, Apostolidou, M, & Takou, E (2012). Working memory and verbal fluency in simultaneous interpreters Working memory and verbal fluency in simultaneous. *Journal of Clinical and Experimental Neuropsychology*, **34**, 37–41.

Tzou, Y-Z, Eslami, ZR, Chen, H-C, & Vaid, J (2012). Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: Evidence from Mandarin-English speakers. *International Journal of Bilingualism*, **16**, 213–227. https://doi.org/10.1177/1367006911403197

Tzou, Y-Z, Eslami, ZR, Chen, H-C, & Vaid, J (2016). Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: Evidence from Mandarin–English speakers. *International Journal of Bilingualism*, **16**, 213–227. https://doi.org/10.1177/1367006911403197

Van de Putte, E, De Baene, W, García-Pentón, L, Woumans, E, Dijkgraaf, A, & Duyck, W (2018). Anatomical and functional changes in the brain after simultaneous interpreting training: A longitudinal study. *Cortex*, **99**, 243–257. https://doi.org/10.1016/j.cortex.2017.11.024

Van Den Noort, M, Bosch, P, Haverkort, M, & Hugdahl, K (2008). A standard computerized version of the reading span test in different languages. *European Journal of Psychological Assessment*, **24**, 35–42. https://doi.org/10.1027/1015-5759.24.1.35

Wen, H, & Dong, Y (2019). How does interpreting experience enhance working memory and short-term memory: a meta-analysis. *Journal of Cognitive Psychology*. https://doi.org/10.1080/20445911.2019.1674857

Yudes, C, Macizo, P, & Bajo, T (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in Psychology*, **2**, 1–9. https://doi.org/10.3389/fpsyg.2011.00309

https://doi.org/10.1017/S1366728920000127 Published online by Cambridge University Press