Bilingualism: Language and Cognition

cambridge.org/bil

Working memory exertion after simultaneous interpreting in bilinguals

Isabelle Chou[1], Agustina Birba[2,3,4], Jiehui Hu[5], Edinson Muñoz[6], Guoqing Kwon[5] and Adolfo M. García[2,6,7]

[1]Sichuan University College of Foreign Languages and Cultures, China; [2]Cognitive Neuroscience Center (CNC), University of San Andres, Argentina; [3]Instituto Universitario de Neurociencia, Universidad de La Laguna, Spain; [4]Facultad de Psicología y Logopedia, Universidad de La Laguna, Spain; [5]School of Foreign Languages, University of Electronic Science and Technology of China, China; [6]Departamento de Lingüística y Literatura, Universidad de Santiago de Chile Facultad de Humanidades, Chile and [7]Global Brain Health Institute (GBHI), University of California, San Francisco, CA, United States & Trinity College Dublin, College Green, Dublin, Ireland

Research Article

Cite this article: Chou, I., Birba, A., Hu, J., Muñoz, E., Kwon, G. and García, A.M. (2026). Working memory exertion after simultaneous interpreting in bilinguals. *Bilingualism: Language and Cognition* 1–11. https://doi.org/10.1017/S1366728925100898

Received: 6 January 2025
Revised: 21 November 2025
Accepted: 25 November 2025

Keywords:
bilingualism; simultaneous interpreting; working memory; cognitive exertion

Corresponding author:
Adolfo M. García,
Email: adolfo.garcia@gbhi.org

Abstract

Simultaneous interpreting (SI), a challenging task enabled by bilingualism, is claimed to distinctly tax working memory (WM). However, causal designs are missing, limiting our understanding of the phenomenon. We recruited 50 Chinese-English bilinguals and assessed their WM performance (alongside inhibitory and fluency outcomes) before and after L1–L2 SI or a control task (text comprehension). WM scores (especially under high-demand, multimodal conditions) increased after text comprehension but not after SI, adjusting for age of L2 appropriation, years of L2 use, L2 proficiency and SI competence. Of note, WM was assessed immediately before and after SI, ruling out other cognitive influences. Conversely, no distinct patterns were observed on inhibitory or fluency tasks. Briefly, this activity seems to hinder practice-related WM gains – a finding that expands contemporary accounts of interpreting.

Highlights

- We examined the impact of simultaneous interpreting (SI) on working memory (WM).
- Unlike a single-language task, SI prevented practice-driven WM boosts.
- This effect emerged for multimodal input, adjusting for bilingual experience factors.
- SI did not have a distinct effect on inhibitory or fluency outcomes.
- Such findings inform cognitive models of bilingualism and SI training guidelines.

1. Introduction

A distinct trait of bilingualism is the possibility to engage in simultaneous interpreting (SI), the immediate rendition of oral messages from one language into another (García et al., 2019). This activity is proposed to greatly tax working memory (WM), but causal designs are wanting, especially including other tasks and cognitive functions (Timarová, 2008). A fertile ground is thus laid to examine the real-time impact of specific bilingual skills on domain-general systems. Here, we tackle this gap via a cognitive exertion paradigm, comparing WM (and other executive) outcomes before and after SI, relative to a control language processing task.

SI ranks among the most demanding cognitive activities enabled by bilingualism (García, 2014; García et al., 2019; Hervais-Adelman et al., 2015). It requires perceiving, accessing and recalling source-language information to reformulate it in a different language as new input segments are processed (Ahrens, 2017; Chernov, 2004; Gerver, 1976; Paradis, 1994, 2009). Such operations occur iteratively, with input typically delivered at >120 words per minute (Chernov, 2004; Gerver, 1975), ear-voice spans lasting 2–4 seconds (Anderson, 1994; Gerver, 1976) and input–output overlaps accounting for 70% of total processing time (Chernov, 1994). These challenges are such that, to prevent task disengagement, reduced output quality and early burnout, professionals are recommended not to perform SI for more than 30 minutes at a time (AAIC, 2025; García, 2019; Moser-Mercer et al., 1998).

Unsurprisingly, the activity entails diverse cognitive costs. SI increases stress responses such as cortisol (Moser-Mercer et al., 1998) and heart rate (Rojo López et al., 2021), especially under heightened time pressure (Korpal, 2016). Also, relative to single-language tasks, like text listening and shadowing, SI involves greater self-reported mental workload (Boos, 2022), objective cognitive exertion markers, including increased pupil dilation (Hyönä et al., 1995; Seeber & Kerzel, 2011), and activation along cortico-subcortical executive brain regions (Hervais-Adelman et al., 2011, 2015). As advanced by several models (Gile, 1999; Mizuno, 2005; Seeber, 2011), SI would place distinct demands on WM, a system supporting the transient

© The Author(s), 2026. Published by Cambridge University Press. 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, provided the original article is properly cited.

storage and combination of information as other processes unfold (Hervais-Adelman & Babcock, 2020).

Indeed, SI entails increased pupil dilation when processing WM-taxing stimuli – for example, asymmetrical structures such as verb-final constructions, as found in Japanese or German (Seeber & Kerzel, 2011). Also, relative to single-language tasks, SI involves higher theta power (a neural marker of WM demands) (Boos, 2022) and prefrontal activation peaks during input–output overlaps (i.e., instances that maximally tax WM) (Hervais-Adelman et al., 2015). Moreover, SI quality correlates with WM outcomes (Bae & Jeong, 2021; Injoque-Ricle et al., 2015; Mellinger & Hanson, 2019), and sustained practice is consistently linked to advantages in this domain (as SI continuously requires *recalling and integrating* input while concurrent processes are deployed) (Chernov, 2004). Such demands seem to peak in the L1–L2 direction, which, relative to the L2–L1 direction, involves greater modulation of WM-relevant brain mechanisms, such as inferior frontal activation (García, 2019; Rinne et al., 2000; Tommola et al., 2000), temporo-frontal connectivity (Zheng, et al., 2020) and frontal theta oscillations (Pérez et al., 2022).

Of note, WM taxation during SI may depend on cognitive demands. Unlike untrained multilinguals, interpreters show training effects in highly taxing (multimodal *n*-back) WM tasks but not in simpler (single-domain) tasks (Morales et al., 2015). Suggestively, no systematic expertise-related advantages are observed in less critically taxed domains, like inhibition (as both languages need to be simultaneously active rather than alternatively suppressed) and verbal fluency (since SI does not require producing arbitrary sequences of decontextualized words) (García et al., 2019).

SI, then, would distinctly tax WM. However, evidence for this claim remains limited. Relevant studies employ cognitively ambiguous markers (e.g., pupil dilation) (Seeber & Kerzel, 2011), WM tests done in the absence of actual SI (Boos, 2022; García et al., 2019) or task-related neural measures failing to isolate or manipulate WM (Boos, 2022; Hervais-Adelman et al., 2015). Moreover, *causal* evidence is scant and limited to comparisons of outcomes before and after months-long training programs. SI practice seems to boost WM performance (Ünlü & Şimşek, 2018) and to lower the engagement of WM-relevant brain regions during the task (Hervais-Adelman et al., 2015), suggesting that this domain is systematically recruited by practicing interpreters. Conversely, other cognitive functions, such as inhibitory skills (Van de Putte et al., 2018), do not exhibit changes before and after SI training, indicating that they are less systematically recruited by this activity. Overall, the real-time impact of SI on WM remains poorly understood, precluding insights into how particular domain-general skills are recruited during specific bilingual processes.

Such gaps can be tackled via cognitive exertion paradigms. In these paradigms, specific domains are assessed before and after a given task, on the assumption that those more critically engaged will be drained and hence impaired upon post-task testing. This approach has revealed critical links between particular tasks and circumscribed cognitive functions, by revealing exertion of inhibitory capacity during emotional suppression (Wessel et al., 2010), learning capacity during extended periods of instruction (Chen et al., 2018), self-regulatory resources during high attentional exertion (Converse & DeShon, 2009) and positive emotion mechanisms during overly enthusiastic behavior (Maranges et al., 2017). More directly, a study on phonological interference effects during SI found that number recall dropped when digit lists were simultaneously translated (as opposed to simply heard or repeated in the same language) before being uttered (Darò & Fabbro, 1994). Similar designs could reveal the causal impact of SI on WM, as the need to transiently store, recall and integrate input information during output production would exhaust relevant mnesic mechanisms and render them suboptimally available for further WM tasks.

Against this background, we conducted a cognitive exertion study on SI. We asked SI trainees to perform a validated WM task (involving unimodal and multimodal stimuli) before and after (a) a 10-minute-long L1–L2 SI session or (b) a control text comprehension (TC) task. To test for specificity, participants also completed tasks tapping other SI-relevant domains (viz., inhibitory control and context-free vocabulary navigation, via Stroop and verbal fluency tasks). We hypothesized that L1–L2 SI would distinctly reduce WM scores, as opposed to less critical mechanisms, such as inhibitory and fluency resources. We further predicted that such an effect would be specific to SI, compared to TC. Finally, we anticipated that the effect would be uninfluenced by inter-individual variability in bilingualism-related variables (age of L2 appropriation,[1] years of use of L2, L2 proficiency, SI competence). Strategically, the use of a high-demand (multimodal integration–recall) and a low-demand (unimodal recall-only) condition allowed testing whether predicted effects were effort-dependent. With this approach, we aim to inform models linking bilingual abilities to specific executive domains.

2. Method

2.1. Participants

We recruited 55 right-handed Chinese-English bilinguals enrolled in the Master's in Translation and Interpreting at the University of Electronic Science and Technology of China (UESTC). They all had normal vision and hearing and reported no neurological impairment. Participation was compensated with course credits. Participants were randomly assigned to the L1–L2 SI or the TC task. Upon removal of five participants due to poor task compliance or withdrawal, the final sample of 50 (Figure 1A) involved 26 SI participants and 24 TC participants – reaching a power of 1 (Supplementary Material 1). Demographic and language profile data for group matching were collected via the Translation and Interpreting Competence Questionnaire (TICQ), a validated (Schaeffer et al., 2020) and previously reported (Chou et al., 2021; Jacob et al., 2024) online self-rating instrument gathering quantitative and qualitative data on demographics, language history and translation/interpreting competence. To favor between-group matching, SI and TC participants were pooled from the same master's courses. TICQ data showed that both groups were matched for sex, age and education, as well as age of L2 appropriation, years of L2 use, L2 proficiency, SI competence and weekly dedication to SI (Table 1). Importantly, all participants showed upper-intermediate levels of L2 proficiency on the TICQ (on a scale of 0–100), consistent with the incidental observation that they each possessed at least one language proficiency certification (Test for English Majors, level 4 or higher; the China Accreditation Test for Translators and Interpreters, level 3; or the International English Language Testing System). All participants signed written

[1]Proposed by Paradis (2009), the term 'appropriation' is a hypernym encompassing both incidental acquisition (which typically occurs before age 7) and metalinguistic learning (more common after age 7). As in previous works (Kogan et al., 2024), this concept is relevant to our design given that we recruited both early and late bilinguals, with no constraints on how they developed their L2 skills.

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

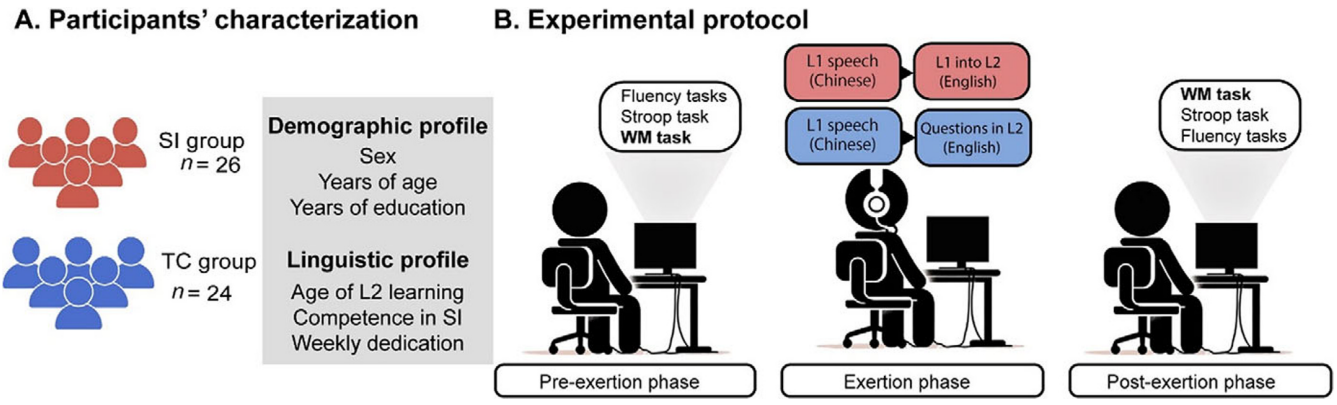


Figure 1. Study design. (A) Fifty Chinese-English bilinguals were randomly assigned to the SI group ($n = 26$) and to the TC group ($n = 24$), both matched for demographic and linguistic variables. (B) The experiment involved a pre-exertion phase (including the WM assessment and control tasks), an exertion phase (involving an SI task for the SI group and a TC task for the TC group) and a post-exertion phase (featuring the same initial tasks in reverse order). SI = simultaneous interpreting, TC = text comprehension.

Table 1. Participants' demographic and linguistic profiles

Demographic profile	SI group	TC group	χ^2/F	*p*-value
Sex (F:M)	22:4	21:3	X^2 (1) = .00	.99
Years of age	23.89 (3.57)	23.42 (1.10)	*F* (1,48) = .38	.53
Years of education	17 (2.44)	18.0 (1.44)	*F*(1,48) = 3.14	.08
Linguistic profile				
Age of L2 appropriation	8.56 (2.71)	8.58 (2.00)	*F*(1,48) = .0017	.96
Years of L2 use	12.1 (4.70)	11.9 (4.62)	*F*(1,48) = .0082	.86
Global L2 proficiency	70.4 (10.2)	73.1 (12.7)	*F*(1,48) = .59	.55
L2 speaking proficiency	66 (9.75)	70.4 (11.2)	*F*(1,48) = 2.26	.14
L2 reading proficiency	76.7 (13.7)	79.4. (14.5)	*F*(1,48) = .47	.50
L2 listening proficiency	69.1 (18.8)	69.3 (14.7)	*F*(1,48) = .0012	.97
L2 writing proficiency	68.3 (11.8)	71.1 (12.7)	*F*(1,48) = .066	.42
L1–L2 SI competence	59.70 (15.44)	66.29 (10.90)	*F*(1,48) = 3.02	.08
L2–L1 SI competence	60.63 (16.89)	67.79 (10.50)	*F*(1,48) = 3.20	.08
Weekly dedication to L1–L2 SI	2.81 (2.83)	5.17 (4.83)	*F*(1,48) = 3.72	.07
Weekly dedication to L2–L1 SI	3.48 (3.35)	4.83 (4.33)	*F*(1,48) = 0.15	.70

Note: Data are shown as mean (*SD*). Statistical comparisons were performed via chi-squared test for sex and via ANOVA for every other variable. SI = simultaneous interpreting, TC = text comprehension.

informed consent. The study protocol was approved by the institutional ethics committee and adhered to the Declaration of Helsinki.

2.2. Study design

The study consisted of three phases (Figure 1B). First, in the pre-exertion phase, participants completed three control tasks (Stroop, semantic fluency, phonemic fluency) and then performed the WM task. Second, in the exertion phase, the SI group performed an L1–L2 SI task, while the TC group completed a TC task. Finally, in the post-exertion phase, participants completed the same tasks of the pre-exertion phase, in reverse order. Across phases, participants sat at a desk in a dimly lit room, facing a desktop computer, with no external distractions. All instructions were delivered orally in L1 (Chinese). Overall, the experiment lasted roughly 2 hours.

2.2.1. Pre-exertion phase

Working memory task. WM was assessed via a widely validated task (Baddeley et al., 2011; Birba et al., 2017; Brockmole et al., 2008) displaying stimulus arrays on a computer screen (Figure 2). Validated in diverse populations (Brockmole et al., 2008; Parra et al., 2009, 2014), this task involves high WM exertion under multimodal integration conditions. Importantly, its non-verbal nature circumvents language-proficiency confounds and priming effects triggered by the SI text. The task included two conditions, tapping on (a) recall and (b) integration–recall skills. In the 'recall-only' condition, participants viewed initial arrays of black shapes and, after a delay, indicated whether a test array featured the same shapes. In the 'integration–recall condition', initial shapes were presented in different colors and, after a delay, participants indicated whether the test array presented the same shape–color

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

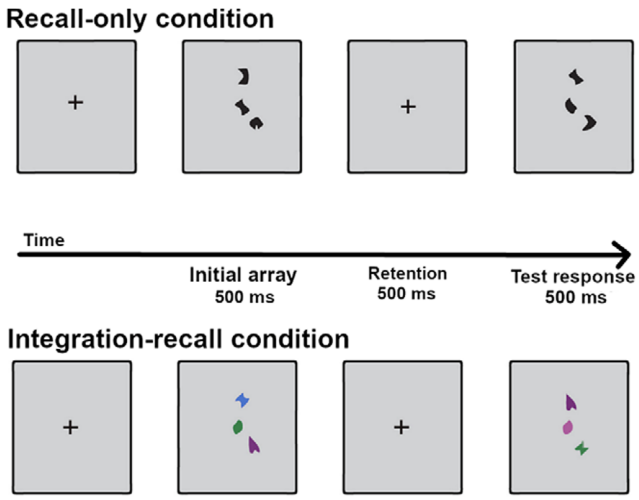


Figure 2. Trial structure in the WM task. The task involves an initial array of items, followed by a retention period and then a test array. Stimuli in the test array may or may not match those of the initial array. The top row illustrates the 'recall-only' condition, tapping on unimodal WM retrieval by manipulating black shapes. The bottom row exemplifies the 'integration–recall' condition, assessing multimodal WM retrieval by manipulating both the shapes and their color.

combinations. A total of 8 six-sided random polygons and eight colors were utilized for the tasks. Stimuli were presented with a standard 17-inch PC monitor with a resolution of 1920*1080 pixels, each occupying 1° of visual angle and positioned within a 10° area. Participants were seated at a viewing distance of 70 cm from the screen.

Each trial commenced with a fixation cross displayed for 500 ms, followed by a random interval lasting between 400 and 600 ms. Initial three-item arrays were then presented for 500 ms, during which participants had to memorize the items. This was followed by a 900-ms retention interval, with no stimuli. A three-item test array appeared next for 500 ms, and participants were required to determine whether these stimuli matched those in the initial display by pressing '1' for 'match' or '2' for 'non-match' (with key assignments counterbalanced across trials). The next trial began only after participants made a response. Half the trials were 'match trials' (with initial and test trials being identical), while the other half were 'non-match trials' (with different initial and test trials).

In the 'recall-only' condition, 'non-match trials' were created by replacing two black shapes from the initial array with new black test shapes, requiring participants to rely solely on shape encoding to detect changes. In the 'integration–recall' condition, 'non-match trials' were created by swapping the colors of two initial items, requiring participants to integrate shape and color information to detect the change. No shape or color was repeated within a single array. Conditions were counterbalanced and trial order was randomized across participants. The task lasts roughly 20 minutes.

Inhibitory control task. Inhibitory control was evaluated via a color–word Stroop task, amply validated in diverse populations, bilinguals with SI experience (Aparicio et al., 2017; MacPherson, 2019; Peng et al., 2017). Facing a computer screen, participants were shown color words ('red', 'yellow', 'blue', 'green') typed in different ink colors, and they had to indicate the name of the ink as quickly as possible by pressing a given key ('f' for red, 'j' for yellow, 'v' for blue, 'n' for green). The task included 16 combinations of words and ink colors, used to create 40 congruent trials (color words written in their corresponding color) and 60 incongruent trials (color words written in a different color), with each congruent combination shown ten times and each incongruent combination shown five times.

Each trial started with a fixation cross displayed for 500 ms, followed by a stimulus word presented for 1,500 ms, during which participants had to respond. A random interval between 500 and 1,000 ms preceded the following trial. All words were presented in the participants' L1. Trials were randomly distributed across congruent and incongruent categories. In the pre-exertion phase, the task began with a practice block, featuring four congruent and 12 incongruent trials. Stimulus administration and response recording were handled via E-Prime 3.0 (Psychology Software Tools, Pittsburgh, PA). Altogether, the task lasted roughly 4 minutes.

Context-free vocabulary navigation tasks. Context-free vocabulary navigation skills were evaluated via one-minute verbal fluency tasks, previously reported in SI research (Dottori et al., 2020; Santilli et al., 2019). The semantic fluency task had participants name as many exemplars of a category as possible, for 60 seconds, first in their L1 and then in their L2. To mitigate practice and priming effects, we counterbalanced the categories 'animals' and 'fruits' between the pre- and post-exertion phases. Within each phase, participants were presented with the same category for both Chinese and English. As to the phonemic fluency task, participants were asked to generate as many words as possible beginning with specific phonemes, orally in English and in written form in Chinese. The order of languages was counterbalanced between the pre- and post-exertion phases, as were the phonemes used for English (/f/, /s/) and for Chinese (/k/, /t/). Following previous studies, we adopted the written modality for Chinese to circumvent issues related to tonal variations of specific characters (e.g., /ma/ means 'mother' with tone 1, 'bother' with tone 2, 'horse' with tone 3, 'curse' with tone 4) (Eng et al., 2019). All responses were audio-recorded in .mp3 format, transcribed by one examiner and subsequently verified by a second examiner to ensure accuracy (a third examiner was summoned to settle the very few instances yielding disagreement). Following validated procedures (Nichols et al., 2019; Selnes, 1991), responses were excluded as invalid if they consisted of either proper names, or repetitions, derivations or affixations of previous responses. In each language, semantic fluency is scored by counting the number of correct, unique words produced within the semantic category in the given time frame. Phonemic fluency scores are calculated by counting the number of correct, non-repeated words in the given time frame.

2.2.2. Exertion phase

The exertion phase involved a specific naturalistic language processing task for each group: SI for the SI group (as our target condition) and TC for the TC group (as our control condition). The stimulus for both tasks consisted of the first 10 minutes of former Singapore Prime Minister Lee Hsien Loong's speech during the 2022 National Day Rally (https://www.pmo.gov.sg/Newsroom/National-Day-Rally-2022-Chinese). The material comprised 1,977 Chinese characters, delivered at a rate of 3.3 characters per second, and it was presented binaurally via a stereo headset after participants adjusted the volume to their preferred level. As in previous works (Bae & Jeong, 2021; Boos et al., 2022; Tzou et al., 2012), no prior information was provided on the topic – a scenario that mirrored participants' training conditions and is often encountered in professional settings (Chang et al., 2018). Note-taking with pen and paper was allowed during the task.

The SI group was instructed to listen to the speech and simultaneously interpret it into English, preserving as much information

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

as possible while producing idiomatic L2 renditions. The use of the L1–L2 direction was strategic to enhance ecological validity, as (i) it is predominant in professional SI across China and (ii) participants were more acquainted with it than with L2–L1 SI in their training program. Output was recorded through the headset's built-in microphone and saved as .mp3 files on Xima 3100 Digital Integrated Teaching System V2.0. SI quality was evaluated according to standard criteria, focusing on content, form and delivery (Hamidi & Pöchhacker, 2007; Macías, 2006; Pöchhacker, 2001). The TC group was asked to listen attentively so as to respond to five open-ended questions presented in a printed survey at the end of the recording. Questions were formulated in L2 (English) and answered in writing. The use of L2 for this task increased task demands (creating more stringent conditions to test the selectivity of predicted SI-related exertion effects), while favoring comparability with the SI task by mirroring its dual-language nature.

2.2.3. Post-exertion phase

The post-exertion phase was identical to the pre-exertion phase, except that tasks were presented in reverse order, with the WM task first, followed by the Stroop and fluency tasks. Through this sequencing, the WM task was always performed immediately before and immediately after the corresponding exertion task, preventing interference from other tests.

2.3. Statistical analysis

Across tasks, accuracy and response time analyses were performed upon removing outlier values at ±2.5 standard deviations from the corresponding group's mean. First, WM task accuracy was analyzed using linear mixed-effects models separately for each 'condition' (recall-only, integration–recall), with 'phase' (pre-exertion, post-depletion) and 'group' (SI, TC) as fixed effects, with a random intercept for subjects. To examine whether group differences were driven by variability in WM performance, we first conducted Levene's tests comparing variances across groups (SI versus TC) and phases (pre-exertion versus post-exertion) for each WM condition (integration–recall and recall-only). In addition, we computed individual change scores (subtraction between the post- and pre-exertion phase outcomes) and visualized their distributions. Group effects on change scores were further assessed using quantile regression at τ = .25, .50 and .75, corresponding to the lower, median and upper parts of the distribution. This approach allowed us to test whether SI selectively influenced participants showing small, typical or large practice-related gains.

Second, for the Stroop task, linear mixed-effects models were separately run on the congruent and incongruent conditions, targeting accuracy and response times (the latter, considering correct responses only). The models included 'phase' (pre-exertion, post-exertion) and 'group' (SI, TC) as fixed effects and 'subject' as a random effect. Additionally, we compared congruent and incongruent trials, for accuracy and response time, using a mixed-effects model with 'condition' (congruent, incongruent), 'phase' (pre-exertion, post-exertion) and 'group' (SI, TC) as fixed effects and 'subject' as a random intercept.

Third, fluency tasks' outcomes were analyzed via linear mixed-effects models, with 'phase' (pre-exertion versus post-exertion), 'language' (L1 versus L2) and 'group' (SI versus TC) as fixed effects, along with their interactions, and a random intercept for 'subjects' to account for repeated measures. Across tasks, ANOVA was used to assess the significance of main effects and interactions. To test for robustness despite inter-individual variability in bilingualism-related factors, all analyses were re-run while covarying for age of L2 appropriation, years of L2 use and L2 proficiency (including relevant macroskills), as well as SI competence in the L1–L2 and L2–L1 directions. Post hoc analyses for significant interactions were conducted via Tukey's HSD tests. Alpha values were set at 0.05. Outlier-adjusted models were also fitted to confirm the robustness of the results. All analyses were performed on R 4.1.1 (Team, 2013).

3. Results

3.1. Working memory

Analysis of accuracy in the integration–recall condition revealed no significant main effects of 'group' ($F_{(1,46)} = 1.89$, $p = .17$) or 'phase' ($F_{(1,46)} = .49$, $p = .48$), but a significant interaction between 'phase' and 'group' ($F_{(1,46)} = 4.8387$, $p = .02$) (Figure 3A). Post hoc comparisons ($MSE = 0.002$, $df = 1, 46$) showed that, relative to the TC group, the SI group performed similarly in the pre-exertion phase ($p = 0.24$) and worse in the post-exertion phase ($p = .02$). This interaction remained significant upon controlling for age of L2 appropriation ($F_{(1,44.89)} = 4.74$, $p = .034$), years of L2 use ($F_{(1,44.50)} = 4.69$, $p = .036$), global L2 proficiency ($F_{(1,44.84)} = 4.75$, $p = .034$), L2 listening proficiency ($F_{(1,44.87)} = 4.76$, $p = .034$), L2 speaking proficiency ($F_{(1,44.79)} = 4.75$, $p = .034$), L2 reading proficiency ($F_{(1,44.89)} = 4.74$, $p = .035$), L1–L2 SI competence ($F_{(1,44.86)} = 4.75$, $p = .035$) and L2–L1 SI competence ($F_{(1,44.86)} = 4.75$, $p = .035$). The same was true for the poorer performance of the SI group compared with the TC group in the post-exertion phase ($p = .02$), upon adjusting for age of L2 appropriation, years of L2 use, L1–L2 SI competence and L2–L1 SI competence. Also, integration–recall accuracy improved from the pre- to the post-exertion phase for the TC group ($M_{pre} = 0.75$ [0.10], $M_{post} = 0.80$ [0.11], $p = .01$), but not for the SI group ($M_{pre} = 0.70$ [0.13], $M_{post} = 0.70$ [0.14], $p = .71$). In the integrated model, with 'condition' (congruent, incongruent) as a fixed effect, we found a significant main effect of 'phase' ($F_{(1,84.68)} = 8.13$, $p = .005$), indicating that participants were more accurate in the second session. There were no other significant effects or interactions, including the three-way interaction ('condition' × 'phase' × 'group': $p = .98$). Response times yielded a significant main effect of 'condition' ($F_{(1,81.22)} = 6.46$, $p = .013$), with slower responses in incongruent compared to congruent trials, and a significant main effect of 'phase' ($F_{(1,82.40)} = 79.15$, $p < .001$), reflecting overall faster responses in the second session. No other effects or interactions were significant (all p-values > .42).

Levene's tests indicated no significant differences in variance between groups or phases for either WM condition (all p-values > .09). Quantile regression analyses of change scores revealed no group differences at lower or median quantiles (τ = .25–.50, all 95% CIs including 0). However, a significant negative effect of SI emerged at the upper quartile (τ = .75: $\beta = -0.05$, 95% CI [−0.087, −0.025]), indicating that the largest practice-related gains observed in controls were attenuated in SI. Complete results and graphical displays are reported in Supplementary Material 2.

3.2. Inhibitory control

Accuracy analyses on the congruent condition revealed no significant main effects or interactions (all p-values > .11). In the incongruent condition, there was a non-significant effect of 'phase' ($F_{(1,28)} = 3.29$, $p = .08$), alongside a non-significant effect of 'group' and a non-significant 'phase-by-group' interaction (both p-values > .83)

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

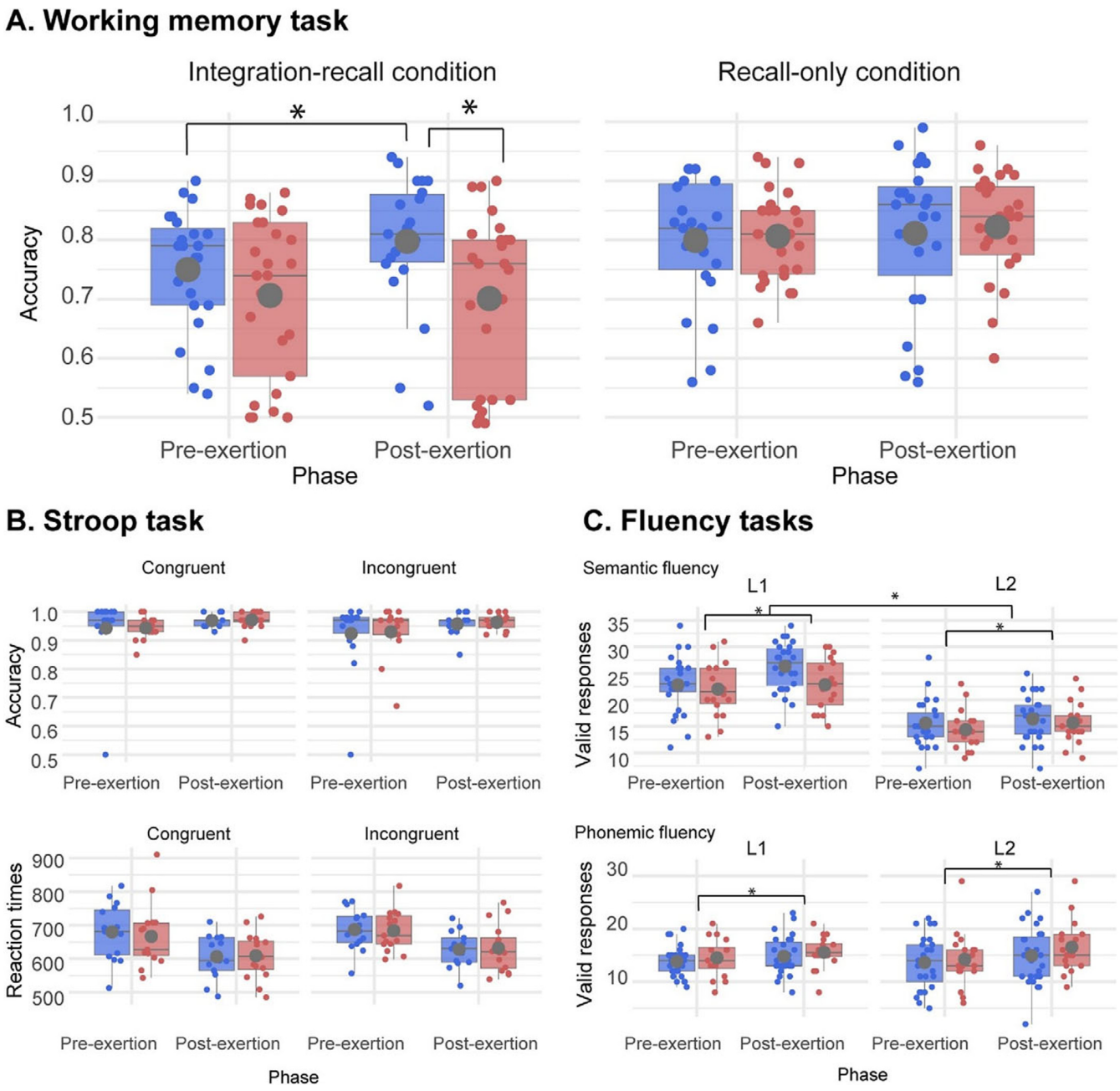


Figure 3. Key results. (A) WM task. Accuracy analyses on the integration–recall condition revealed similar performance between groups in the pre-exertion phase and worse performance for the SI than for the TC group in the post-exertion phase. Also, performance improved between sessions for the TC group but not for the SI group. No significant accuracy effects emerged in the recall-only condition. (B) Inhibitory control task. No significant differences were found between groups or phases for either accuracy or response time. (C) Context-free vocabulary navigation tasks. Semantic (top inset) and phonemic (bottom inset) fluency results revealed significant main effects of phase, with significant increases from pre- to post-exertion. Red and blue dots represent SI and TC group participants, respectively. SI = simultaneous interpreting, TC = text comprehension.

(Figure 3B, top inset). Response time analyses on the congruent condition showed a non-significant effect of 'phase' ($F_{(1,\ 28)} = 3.47$, $p = .08$), together with a non-significant effect of 'group' and a non-significant 'phase-by-group' interaction (both p-values > .66). The incongruent condition revealed no significant effect (all p-values > .30) (Figure 3B, bottom inset). In the integrated model, with 'condition' (congruent, incongruent) as a fixed effect, we found a significant main effect of phase ($F_{(1,\ 84.68)} = 8.13$, $p = .005$), indicating that participants were more accurate in the second session. There were no other significant effects or interactions, including the three-way interaction ('condition' × 'phase' × 'group': $p = .98$). For response time, the model revealed a significant main effect of condition ($F_{(1,\ 81.22)} = 6.46$, $p = .013$), with slower responses in incongruent compared to congruent trials, and a significant main effect of phase ($F_{(1,\ 82.40)} = 79.15$, $p < .001$), reflecting overall faster responses in the second session. No other effects or interactions were significant (all p-values > .42).

3.3. Context-free vocabulary navigation

Semantic fluency results (Figure 3C, top inset) revealed significant main effects of 'language' ($F_{(1,126)} = 191.56$, $p < .001$) and 'phase'

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

($F_{(1,126)} = 7.13$, $p = .009$), the latter revealing a significant increase in semantic fluency from the pre- to the post-exertion phase. Phonemic fluency outcomes (Figure 3C, bottom inset) showed a significant main effect of 'phase' ($F_{(1,126)} = 9.13$, $p = .003$), typified by a significant increase in phonemic fluency performance from the pre- to the post-exertion phase. No other main or interaction effects reached significance in either fluency task (all p-values > .24).

4. Discussion

This study investigated the impact of SI on WM via a cognitive exertion paradigm. We found that SI, as opposed to TC, prevented practice-driven boosting of integration–recall outcomes, even adjusting for key aspects of bilingual experience. No such effects were observed on inhibitory and fluency performance. These results shed light on the cognitive demands of SI as a distinct form of bilingual processing.

Prior works have shown that WM outcomes are associated with SI experience (García et al., 2019) as well as interpreting quality (Bae & Jeong 2021; Injoque-Ricle et al., 2015; Mellinger & Hanson, 2019), and relevant electrophysiological signatures are distinctly modulated during SI (Boos, 2022). Our pre/post study extends such findings, indicating that WM resources are markedly affected after 10 minutes of this activity.

Crucially, post-task WM performance differed between groups. Integration–recall outcomes improved after TC. This likely reflects well-established practice and repetition effects in WM tasks, manifested as measurable gains through increased familiarity, greater encoding efficiency or more strategic resource allocation (Chein & Morrison, 2010; Jonides et al., 2008). Conversely, no such improvement emerged after SI, suggesting that this task hindered the consolidation or reactivation of learning-related mechanisms. This aligns with resource depletion models, which posit that high-load tasks temporarily exhaust executive resources, impairing subsequent performance or learning (Chen & Kalyuga, 2020; Just & Carpenter, 1992).

In the present design, the selective absence of WM gains after SI could reflect a combination of (1) acute resource depletion during SI and (2) interference with consolidation processes due to the multitasking demands of interpreting (Christoffels et al., 2006). Additionally, the greater variability observed in WM performance within the SI group may reflect inter-individual differences in cognitive resilience or interpreting strategies, further supporting the view that SI imposes heterogeneous cognitive loads that differentially affect WM outcomes (Dong et al., 2018). Importantly, these group differences were not explained by increased variability: variance did not differ between groups or phases, and quantile analyses indicated that only the control group showed practice-related gains at the upper tail of the integration–recall distribution. Thus, SI did not enhance heterogeneity per se, but rather prevented the emergence of high-gain learners typically observed after practice.

Again, such a disruption emerged only in the integration–recall condition, which, unlike the recall-only condition, requires binding different processes (viz., shape and color recognition) (Baddeley et al., 2011; Birba et al., 2017; Brockmole et al., 2008). Suggestively, and as corroborated in the TC group, integration–recall mechanisms are more susceptible to training and interference effects than recall-only mechanisms (Allen et al., 2006; Parra et al., 2009) – a finding that highlights the selectivity of SI-related WM effects for high-load conditions. Compatibly, cognitive advantages associated with sustained SI practice are systematic for WM tasks that involve concurrent cognitive operations but inconsistent for tasks that tap on unimodal retrieval (García et al., 2019). For example, interpreters show training effects in WM tasks involving both audio and visual stimuli, but not under a single-modality condition (Morales et al., 2015). This selectivity might reflect the critical role of resource management and task scheduling during SI, two aspects that are distinctly taxed during multimodal WM processes (Mizuno, 2005).

The above pattern survived covariation with age of L2 appropriation, years of L2 use, L2 proficiency and SI competence. This is noteworthy, as WM performance in bilinguals is often influenced by such factors (Köpke & Nespoulous, 2006; Lukasik et al., 2018; Miyake & Friedman, 2013; Tzou et al., 2012). Thus, contrary to other WM dynamics in this population, SI-driven WM exertion seems consistent beyond inter-individual variability in bilingual experience. Moreover, such WM exertion effect was not observed in the TC group, mirroring evidence that other SI-related effects on executive mechanisms are absent in single-language tasks, like text listening and shadowing (Boos, 2022; Hervais-Adelman et al., 2011, 2015; Hyönä et al., 1995). Since the WM task was identical (in stimuli, structure and duration) for both groups, the observed draining of WM resources would be a differential consequence of SI rather than a manifestation of sustained discourse processing (or prior WM taxing) at large.

Furthermore, exertion effects during L1–L2 SI seem distinctly operative in WM functions, as they were not observed on inhibitory or fluency outcomes. Similarly, while SI experience is consistently associated with WM advantages, no systematic enhancement is observed on either of those domains (Aparicio et al., 2017; García et al., 2019; Morales et al., 2015; Stavrakaki et al., 2012; Van de Putte et al., 2018). For example, despite inducing neuroanatomical changes, SI training yielded null effects on a Simon task – a gold-standard measure of response inhibition (Van de Putte et al., 2018). As proposed elsewhere (García et al., 2019), this might be so because concomitant processes during SI (e.g., source-language comprehension and target-language production) need to be active in parallel rather than alternately suppressed. Still, null effects on inhibition and fluency may be partly driven by our design. Indeed, both tests, unlike the WM task, had demands other than SI between their pre- and post-exertion versions. Though further research is needed to rule out this confound, our findings suggest that SI does not uniformly impact domain-general processes at large.

Incidentally, fluency results revealed better outcomes in L1 (Chinese) than in L2 (English). This pattern might reflect two phenomena. First, in Asian languages, animal fluency outcomes evince disproportionately high rates of zodiac animals, a list that is rotely learned and quickly produced across L1 speakers (Sung et al., 2025). Second, word retrieval outcomes are generally lower in L2 than in L1, likely due to less entrenchment in the former (French & Jacquet, 2004) – a difference that may override immediate L1–L2 priming effects. However, fluency results might be influenced by the fixed administration of the L1 before L2 condition, inviting further research with strategic manipulations.

Our study carries theoretical implications. Links between SI and WM are acknowledged in different accounts, but mainly supported by cross-sectional studies (García et al., 2019). Rooted in a pre/post design, our results suggest that SI may be causally involved in the exhaustion of WM resources, offering mechanistic constraints for cognitive modeling. More particularly, our findings support the view that SI effects are typified by 'demand-based domain-specificity' (García et al., 2019: 736). While this claim was first advanced by reference to SI-induced benefits (García, 2014), the abolition of practice-related WM effects shows that even SI-induced disadvantages are also confined to functions distinctly

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

recruited during the task. This extends the interpreter advantage hypothesis, suggesting that the selective impact of SI on its central domains may manifest not only as long-term gains but also as immediate decline. Accordingly, accounts of SI-related effects could be fruitfully expanded by detailing how immediate resource depletion progressively leads to diachronic boosts – e.g., through entrenchment or activation threshold adaptations.

This work is not without limitations. First, though adequately powered and similar to that of previous works (e.g., Babcock & Vallesi, 2017; Becker et al., 2016; Čeňková et al., 2014; Tzou et al., 2012), our sample size was modest. As in previous bilingualism studies (e.g., Bialystok et al., 2004; Calabria et al., 2012; Blanco-Elorrieta & Pylkkänen, 2017; Declerck et al., 2015), this was mitigated via linear mixed-effects models with a simple structure and within-subject data, which enhances statistical sensitivity by reducing inter-individual variability (Charness et al., 2012). Nonetheless, future studies should aim to replicate these findings with larger and more diverse samples. Second, the sample consisted solely of interpreting students. Given that cognitive outcomes may vary widely between trainees and professionals (García et al., 2019), future works should test whether similar exertion effects occur across both populations. Third, L2 proficiency and SI competence levels were based solely on self-report data. Though widely used (Hulstijn, 2012) and predictive of objective language outcomes (Langdon et al., 2005; Marian et al., 2007; Gollan et al., 2012), these could be prone to self-image and desirability biases. Future works should complement such measures with formal test results as inclusion criteria. Fourth, participants received no topic information prior to SI. While this mirrors previous works (Bae & Jeong, 2021; Boos et al. 2022; Tzou et al. 2012), frequent professional scenarios (Chang et al. 2018) and participants' experience during formal training, it deviates from standard, recommended practice, inviting studies on how pre-task briefing might influence WM demands.

Fifth, while our WM task has been explicitly used to operationalize recall skills (Brockmole et al., 2008), these are variously intertwined with recognition abilities, especially in the present paradigm. Future works could tackle this issue through tasks that more clearly demarcate both domains, disentangling the possible contributions of recognition skills to present results. Sixth, although the use of colors as stimuli minimizes semantic associations, their nameability raises the possibility of verbal coding strategies (Baddeley, 2003). While our data do not suggest that participants systematically engaged in such strategies, this possibility cannot be fully excluded, and thus, present effects should be interpreted as reflecting domain-general, multimodal and possibly language-linked mechanisms. Seventh, the SI session may have been too short for exertion effects to manifest in the post-SI inhibitory and fluency measures. Future works could examine whether stronger or longer depletion influences tasks administered later in the protocol. Eighth, inhibition and fluency tasks were administered after the post-training WM task. As such, their underlying resources may have been depleted but then replenished by the time of testing. While the short delay (20 minutes) and the added fatigue from the WM task make full recovery unlikely, future studies should counterbalance tasks to more firmly assess the selectivity of observed WM effects. Alternatively, SI-induced exertion effects may not have been strong enough to persist until the inhibitory and fluency tasks were performed. Further studies could assess this possibility with a prolonged SI session. Ninth, while our pre/post design offers an advantage over purely cross-sectional approaches by capturing within-subject change (Charness et al. 2012), our main finding – the interaction between group and phase – relies on a between-subjects comparison. Observed differences, then, could be partly influenced by confounds beyond our design. Within-subject or crossover designs would help to more precisely isolate the phenomena targeted herein.

Finally, our SI task involved the L1–L2 direction only. Though widespread in the Chinese context, this direction involves neurocognitive demands that differ from those of L2–L1 interpreting (Pérez et al., 2022; Rinne et al., 2000; Zheng et al., 2020), as evidenced by distinct behavioral performance and greater activation along inferior frontal regions implicated in WM and linguistic processes (García, 2019; Rinne et al., 2000; Tommola et al., 2000). Higher demands for the L1–L2 direction have also been reported during sentence and word translation, manifested as increased temporo-frontoparietal connectivity (Zheng, et al., 2020) and higher frontal theta power (Pérez et al., 2022), respectively. Therefore, present results should be taken to reflect L1–L2 SI demands in particular, calling for further studies on how directionality modulates WM exertion.

5. Conclusion

In sum, our study shows that a brief period of SI interferes with practice-driven boosting of WM integration–recall outcomes. This effect seems specific to both SI (relative to single-language discourse processing) and WM (as opposed to other cognitive domains). Future works along this line can further illuminate the interplay between cross-linguistic tasks and the cognitive mechanisms that support performance in bilingual populations.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S1366728925100898.

Data availability statement. All experimental data and scripts are available via the Open Science Framework at https://osf.io/p37wr/?view_only=73e68d4e0e254400a611f9cb94d06c90.

Author contribution. Isabelle Chou: data curation, methodology, software, formal analysis, project administration and writing – original draft. Agustina Birba: methodology, software, formal analysis and writing – original draft. Jiehui Hu: resources, funding acquisition, investigation and writing – review and editing. Edinson Muñoz: investigation and writing – review and editing. Guoqing Kwon: data curation, methodology, software and writing – review and editing. Adolfo M. García: conceptualization, methodology, supervision and writing – review and editing.

Funding statement. Adolfo García is partially supported by the National Institute on Aging of the National Institutes of Health (R01AG075775); ANID (FONDECYT Regular 1250317, 1250091); Universidad de Santiago de Chile (DICYT 032351MA); and the Multi-partner Consortium to Expand Dementia Research in Latin America (ReDLat), which is supported by the Fogarty International Center and the National Institutes of Health, the National Institute on Aging (R01AG057234, R01AG075775, R01AG21051 and CARDS-NIH), Alzheimer's Association (SG-20-725707), Rainwater Charitable Foundation's Tau Consortium, the Bluefield Project to Cure Frontotemporal Dementia and the Global Brain Health Institute. This study was supported by the National Foreign Expert Project of China (grant no. G2022167001L). The contents of this publication are solely the responsibility of the authors and do not represent the official views of these institutions.

Competing interests. The authors declare none.

References

AAIC (2025). Frequently asked questions. https://aiic.ch/faq/#why-dont-simultaneous-interpreters-work-alone

Ahrens, B. (2017). Interpretation and cognition. In J. W. Schwieter & A. Ferreira (Eds.), *The handbook of translation and cognition* (pp. 445–460). John Wiley & Sons.

Allen, R. J., **Baddeley, A. D.**, & **Hitch, G. J.** (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, **135**(2), 298–313. https://doi.org/10.1037/0096-3445.135.2.298.

Anderson, L. (1994). Simultaneous interpretation: Contextual and translational aspects. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 101–120). John Benjamins.

Aparicio, X., **Heidlmayr, K.**, & **Isel, F.** (2017). Inhibition efficiency in highly proficient bilinguals and simultaneous interpreters: Evidence from language switching and Stroop tasks. *Journal of Psycholinguistic Research*, **46**, 1427–1451. https://doi.org/10.1007/s10936-017-9501-3.

Babcock, L., & **Vallesi, A.** (2017). Are simultaneous interpreters expert bilinguals, unique bilinguals, or both? *Bilingualism: Language and Cognition*, **20**(2), 403–417.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, **4**(10), 829–839.

Baddeley, A. D., **Allen, R. J.**, & **Hitch, G. J.** (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, **49**(6), 393–400. https://doi.org/10.1016/j.neuropsychologia.2010.12.042.

Bae, M., & **Jeong, C. J.** (2021). The role of working memory capacity in interpreting performance: An exploratory study with student interpreters. *Translation, Cognition & Behavior*, **4**(1), 26–46. https://doi.org/10.1075/tcb.00050.bae.

Becker, M., **Schubert, T.**, **Strobach, T.**, **Gallinat, J.**, & **Kühn, S.** (2016). Simultaneous interpreters vs. professional multilingual controls: Group differences in cognitive control as well as brain structure and function. *NeuroImage*, **134**, 250–260. https://doi.org/10.1016/.neuroimage.2016.03.079.

Bialystok, E., **Craik, F. I.**, **Klein, R.**, & **Viswanathan, M.** (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, **19**(2), 290.

Birba, A., **Hesse, E.**, **Sedeño, L.**, **Mikulan, E. P.**, **García, M. d. C.**, **Ávalos, J.**, **Adolfi, F.**, **Legaz, A.**, **Bekinschtein, T. A.**, & **Zimerman, M.** (2017). Enhanced working memory binding by direct electrical stimulation of the parietal cortex. *Frontiers in Aging Neuroscience*, **9**, 178. https://doi.org/10.3389/fnagi.2017.00178.

Blanco-Elorrieta, E., & **Pylkkänen, L.** (2017). Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *Journal of Neuroscience*, **37**(37), 9022–9036.

Boos, M., **Kobi, M.**, **Elmer, S.**, & **Jäncke, L.** (2022). The influence of experience on cognitive load during simultaneous interpretation. *Brain and Language*, **234**, 105–119. https://doi.org/10.1016/j.bandl.2022.105185.

Brockmole, J. R., **Parra, M. A.**, **Sala, S. D.**, & **Logie, R. H.** (2008). Do binding deficits account for age-related decline in visual working memory? *Psychonomic Bulletin & Review*, **15**, 543–547. https://doi.org/10.3758/pbr.15.3.543.

Calabria, M., **Hernández, M.**, **Branzi, F. M.**, & **Costa, A.** (2012). Qualitative differences between bilingual language control and executive control: Evidence from task-switching. *Frontiers in Psychology*, **2**, 399.

Čeňková, I., **Meylaerts, R.**, **Hertog, E.**, **Szmalec, A.**, & **Duyck, W.** (2014). Simultaneous interpreting and working memory executive control. *Interpreting*, **16**(2), 139–168. https://doi.org/10.1075/intp.16.2.01tim.

Chang, C. C., **Wu, M. M. C.**, & **Kuo, T. C. G.** (2018). Conference interpreting and knowledge acquisition: How professional interpreters tackle unfamiliar topics. *Interpreting*, **20**(2), 204–231. https://doi.org/10.1075/intp.00010.cha.

Charness, G., **Gneezy, U.**, & **Kuhn, M. A.** (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, **81**(1), 1–8. https://doi.org/10.1016/j.jebo.2011.08.009.

Chein, J. M., & **Morrison, A. B.** (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review*, **17**(2), 193–199. https://doi.org/10.3758/PBR.17.2.193.

Chen, O., **Castro-Alonso, J. C.**, **Paas, F.**, & **Sweller, J.** (2018). Extending cognitive load theory to incorporate working memory resource exertion: Evidence from the spacing effect. *Educational Psychology Review*, **30**, 483–501. https://doi.org/10.1007/s10648-017-9426-2.

Chen, O., & **Kalyuga, S.** (2020). Cognitive load theory: When, why, and how. In J. Dunlosky & K. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 87–116). Cambridge University Press.

Chernov, G. V. (1994). Message redundancy and message anticipation in simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 139–153). John Benjamins.

Chernov, G. V. (2004). *Inference and anticipation in simultaneous interpreting: A probability-prediction model*. John Benjamins.

Chou, I., **Hu, J.**, **Muñoz, E.**, & **García, A. M.** (2021). Discourse-level information recall in early and late bilinguals: Evidence from single-language and cross-linguistic tasks. *Frontiers in Psychology*, **12**, 757351. https://doi.org/10.3389/fpsyg.2021.757351.

Christoffels, I. K., **de Groot, A. M. B.**, & **Kroll, J. F.** (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, **54**(3), 324–345. https://doi.org/10.1016/j.jml.2005.12.004.

Converse, P. D., & **DeShon, R. P.** (2009). A tale of two tasks: Reversing the self-regulatory resource exertion effect. *Journal of Applied Psychology*, **94**(5), 1318–1324. https://doi.org/10.1037/a0014604.

Darò, V., & **Fabbro, F.** (1994). Verbal memory during simultaneous interpretation: Effects of phonological interference. *Applied Linguistics*, **15**(4), 365–381. https://doi.org/10.1093/applin/15.4.365.

Declerck, M., **Thoma, A. M.**, **Koch, I.**, & **Philipp, A. M.** (2015). Highly proficient bilinguals implement inhibition: Evidence from n-2 language repetition costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **41**(6), 1911.

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**, 875. https://doi.org/10.3389/fpsyg.2018.00875.

Dottori, M., **Hesse, E.**, **Santilli, M.**, **Vilas, M. G.**, **Caro, M. M.**, **Fraiman, D.**, **Sedeño, L.**, **Ibáñez, A.**, & **García, A. M.** (2020). Task-specific signatures in the expert brain: Differential correlates of translation and reading in professional interpreters. *NeuroImage*, **209**, 116519. https://doi.org/10.1016/j.neuroimage.2020.116519.

Eng, N., **Vonk, J. M.**, **Salzberger, M.**, & **Yoo, N.** (2019). A cross-linguistic comparison of category and letter fluency: Mandarin and English. *Quarterly Journal of Experimental Psychology*, **72**(3), 651–660. https://doi.org/10.1177/1747021818765997.

French, R., & **Jacquet, M.** (2004). Understanding bilingual memory: Models and data. *Trends in Cognitive Science*, **8**(2), 87–93.

García, A. M. (2014). The interpreter advantage hypothesis preliminary data patterns and empirically motivated questions. *Translation and Interpreting Studies*, **9**(2), 219–238. https://doi.org/10.1075/tis.9.2.04gar.

García, A. M. (2019). *The neurocognition of translation and interpreting*. John Benjamins.

García, A. M., **Muñoz, E.**, & **Kogan, B.** (2019). Taxing the bilingual mind: Effects of simultaneous interpreting experience on verbal and executive mechanisms. *Bilingualism: Language and Cognition*, **23**(4), 1–11. https://doi.org/10.1017/S1366728919000063.

Gerver, D. (1975). A psychological approach to simultaneous interpretation. *Meta*, **20**(2), 119–128.

Gerver, D. (1976). Empirical studies of simultaneous interpretation: A review and a model. In R. W. Brislin (Ed.), *Translation: Applications and research* (pp. 165–207). Gardner Press.

Gile, D. (1999). Testing the effort models' tightrope hypothesis in simultaneous interpreting-a contributionv. *HERMES-Journal of Language and Communication in Business*, **23**, 153–172. https://doi.org/10.7146/hjlcb.v12i23.25553.

Gollan, T. H., **Weissberger, G. H.**, **Runnqvist, E.**, **Montoya, R. I.**, & **Cera, C. M.** (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, **15**(3), 594–615.

Hamidi, M., & **Pöchhacker, F.** (2007). Simultaneous consecutive interpreting: A new technique put to the test. *Meta*, **52**(2), 276–289. https://doi.org/10.7202/016070ar.

Hervais-Adelman, A., & **Babcock, L.** (2020). The neurobiology of simultaneous interpreting: Where extreme language control and cognitive control intersect. *Bilingualism: Language and Cognition*, **23**(4), 740–751. https://doi.org/10.1017/S1366728919000324.

Hervais-Adelman, A., **Moser-Mercer, B.**, & **Golestani, N.** (2011). Executive control of language in the bilingual brain: Integrating the evidence from

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

neuroimaging to neuropsychology [review]. *Frontiers in Psychology*, **2**, 234. https://doi.org/10.3389/fpsyg.2011.00234.

Hervais-Adelman, A., **Moser-Mercer, B.**, **Michel, C. M.**, & **Golestani, N.** (2015). fMRI of simultaneous interpretation reveals the neural basis of extreme language control. *Cerebral Cortex*, **25**(12), 4727–4739. https://doi.org/10.1093/cercor/bhu158.

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition* **15**(2), 422–433.

Hyönä, J., **Tommola, J.**, & **Alaja, A. M.** (1995). Pupil dilation as a measure of processing load in simultaneous interpreting and other language tasks. *The Quarterly Journal of Experimental Psychology*, **48**(3), 598–612. https://doi.org/10.1080/14640749508401407

Injoque-Ricle, I., **Barreyro, J. P.**, **Formoso, J.**, & **Jaichenco, V. I.** (2015). Expertise, working memory and articulatory suppression effect: Their relation with simultaneous interpreting performance. *Advances in Cognitive Psychology*, **11**(2), 56–63. https://doi.org/10.5709/acp-0171-1.

Jacob, G., **Schaeffer, M. J.**, **Oster, K.**, & **Hansen-Schirra, S.** (2024). The psycholinguistics of shining-through effects in translation: Cross-linguistic structural priming or serial lexical co-activation? *Applied PsychoLinguistics*, **45**(3), 542–566. https://doi.org/10.1017/S0142716424000183.

Jonides, J., **Lewis, R. L.**, **Nee, D. E.**, **Lustig, C. A.**, **Berman, M. G.**, & **Moore, K. S.** (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, **59**, 193–224. https://doi.org/10.1146/annurev.psych.59.103006.093615.

Just, M. A., & **Carpenter, P. A.** (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**(1), 122–149.

Kogan, B., **Agulla, L.**, **Dottori, M.**, **Amoruso, L.**, **Vivas, L.**, & **García, A. M.** (2024). Do we mean the same? Semantic native-likeness in highly proficient second-language users. *International Journal of Bilingualism*. https://doi.org/10.1177/13670069241267136.

Köpke, B., & **Nespoulous, J. L.** (2006). Working memory performance in expert and novice interpreters. *Interpreting*, **8**(1), 1–23. https://doi.org/10.1075/intp.8.1.02kop.

Korpal, P. (2016). Interpreting as a stressful activity: Physiological measures of stress in simultaneous interpreting. *Poznan Studies in Contemporary Linguistics*, **52**(2), 297–316. https://doi.org/10.1515/psicl-2016-0011.

Langdon, H. W., **Wiig, E. H.**, & **Nielsen, N. P.** (2005). Dual-dimension naming speed and language-dominance ratings by bilingual Hispanic adults. *Bilingual Research Journal*, **29**(2), 319–336.

Lukasik, K. M., **Lehtonen, M.**, **Soveri, A.**, **Waris, O.**, **Jylkkä, J.**, & **Laine, M.** (2018). Bilingualism and working memory performance: Evidence from a large-scale online study. *PLoS One*, **13**(11), e0205916. https://doi.org/10.1371/journal.pone.0205916.

Macías, E. M. P. (2006). Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency. *Interpreting*, **8**(1), 25–43. https://doi.org/10.1075/intp.8.1.03pra.

MacPherson, M. K. (2019). Cognitive load affects speech motor performance differently in older and younger adults. *Journal of Speech, Language, and Hearing Research*, **62**(5), 1258–1277. https://doi.org/10.1044/2018_JSLHR-S-17-0222.

Maranges, H. M., **Brandon, J.**, **Schmeichel, B. J.**, & **Baumeister, R. F.** (2017). Comparing cognitive load and self-regulatory depletion: Effects on emotions and cognitions. *Learning and Instruction*, **51**, 74–84. https://doi.org/10.1016/j.learninstruc.2016.10.010.

Marian, V., **Blumenfeld, H. K.**, & **Kaushanskaya, M.** (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Language*.

Mellinger, C. D., & **Hanson, T. A.** (2019). Meta-analyses of simultaneous interpreting and working memory. *Interpreting*, **21**(2), 165–195. https://doi.org/10.1075/intp.00026.mel.

Miyake, A., & **Friedman, N. P.** (2013). Individual differences in second language proficiency: Working memory as language aptitude. In Alice F. Healy, Lyle E. Bourne, Jr. (Ed.), *Foreign language learning* (pp. 339–364). Psychology Press.

Mizuno, A. (2005). Process model for simultaneous interpreting and working memory. *Meta*, **50**(2), 739–752. https://doi.org/10.7202/011015ar.

Morales, J., **Padilla, F.**, **Gómez-Ariza, C. J.**, & **Bajo, M. T.** (2015). Simultaneous interpretation selectively influences working memory and attentional networks. *Acta Psychologica*, **155**, 82–91. https://doi.org/10.1016/j.actpsy.2014.12.004.

Moser-Mercer, B., **Künzli, A.**, & **Korac, M.** (1998). Prolonged turns in interpreting: Effects on quality, physiological and psychological stress (pilot study). *Interpreting*, **3**(1), 47–64. https://doi.org/10.1075/intp.3.1.03mos.

Nichols, E., **Szoeke, C. E. I.**, **Vollset, S. E.**, **Abbasi, N.**, **Abd-Allah, F.**, **Abdela, J.**, **Aichour, M. T. E.**, **Akinyemi, R. O.**, **Alahdab, F.**, **Asgedom, S. W.**, **Awasthi, A.**, **Barker-Collo, S. L.**, **Baune, B. T.**, **Béjot, Y.**, **Belachew, A. B.**, **Bennett, D. A.**, **Biadgo, B.**, **Bijani, A.**, **Bin Sayeed, M. S.**, & **Murray, C. J. L.** (2019). Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology*, **18**(1), 88–106. https://doi.org/10.1016/S1474-4422(18)30403-4.

Paradis, M. (1994). Towards a neurolinguistic theory of simultaneous translation: The framework. *International Journal of PsychoLinguistics*, **10**, 319–335.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins.

Parra, M. A., **Abrahams, S.**, **Logie, R. H.**, **Méndez, L. G.**, **Lopera, F.**, & **Della Sala, S.** (2009). Visual short-term memory binding deficits in familial Alzheimer's disease. *Brain*, **132**(4), 1099–1110. https://doi.org/10.1093/brain/awq148.

Parra, M. A., **Della Sala, S.**, **Logie, R. H.**, & **Morcom, A. M.** (2014). Neural correlates of shape–color binding in visual working memory. *Neuropsychologia*, **52**, 27–36. https://10.1016/j.neuropsychologia.2013.09.036.

Peng, H., **Gao, Y.**, & **Mao, X.** (2017). The roles of sensory function and cognitive load in age differences in inhibition: Evidence from the Stroop task. *Psychology and Aging*, **32**(1), 42–50. https://doi.org/10.1037/pag0000149.

Pérez, G., **Hesse, E.**, **Dottori, M.**, **Birba, A.**, **Amoruso, L.**, **Martorell Caro, M.**, **Ibáñez, A.**, & **García, A. M.** (2022). The bilingual lexicon, back and forth: Electrophysiological signatures of translation asymmetry. *Neuroscience*, **481**, 134–143. https://doi.org/10.1016/j.neuroscience.2021.11.046.

Pöchhacker, F. (2001). Quality assessment in conference and community interpreting. *Meta*, **46**(2), 410–425. https://doi.org/10.7202/003847ar.

Rinne, J. O., **Tommola, J.**, **Laine, M.**, **Krause, B. J.**, **Schmidt, D.**, **Kaasinen, V.**, & **Sunnari, M.** (2000). The translating brain: Cerebral activation patterns during simultaneous interpreting. *Neuroscience Letters*, **294**(2), 85–88. https://doi.org/10.1016/S0304-3940(00)01540-8.

Rojo López, A., **Foulquié, A.**, **López, L. E.**, & **Martínez Sánchez, F.** (2021). Analysis of speech rhythm and heart rate as indicators of stress on student interpreters. *Perspectives: Studies in Translation Theory and Practice*, **29**(4), 591–607. https://doi.org/10.1080/0907676x.2021.1900305.

Santilli, M., **Vilas, M. G.**, **Mikulan, E.**, **Caro, M. M.**, **Muñoz, E.**, **Sedeño, L.**, **Ibáñez, A.**, & **García, A. M.** (2019). Bilingual memory, to the extreme: Lexical processing in simultaneous interpreters. *Bilingualism: Language and Cognition*, **22**(2), 331–348. https://doi.org/10.1017/S1366728918000378.

Schaeffer, M., **Huepe, D.**, **Hansen-Schirra, S.**, **Hofmann, S.**, **Muñoz, E.**, **Kogan, B.**, **Herrera, E.**, **Ibáñez, A.**, & **García, A. M.** (2020). The translation and interpreting competence questionnaire: An online tool for research on translators and interpreters. *Perspectives: Studies in Translation Theory and Practice*, **28**(1), 90–108. https://doi.org/10.1080/0907676X.2019.1629468.

Seeber, K. G. (2011). Cognitive load in simultaneous interpreting: Existing theories—New models. *Interpreting*, **13**(2), 176–204. https://doi.org/10.1177/1367006911402982.

Seeber, K. G., & **Kerzel, D.** (2011). Cognitive load in simultaneous interpreting: Model meets data. *International Journal of Bilingualism*, **16**(2), 228–242. https://doi.org/10.1177/1367006911402982.

Selnes, O. A. (1991). A compendium of neuropsychological tests: Administration, norms, and commentary. *Neurology*, **41**(11), 1856–1856-a. https://doi.org/10.1212/WNL.41.11.1856-a.

Stavrakaki, S., **Megari, K.**, **Kosmidis, M. H.**, **Apostolidou, M.**, & **Takou, E.** (2012). Working memory and verbal fluency in simultaneous interpreters. *Journal of Clinical and Experimental Neuropsychology*, **34**(6), 624–633. https://doi.org/10.1080/13803395.2012.667068.

Sung, J. E., **Shin, J.**, **Scimeca, M.**, **Li, R.**, & **Kiran, S.** (2025). Cross-linguistic and multicultural effects on animal fluency performance in persons with aphasia. *American Journal of Speech-Language Pathology*, 1–11Advance online publication. https://doi.org/10.1044/2025_AJSLP-24-00398.

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press

Team, R. C. (2013). *R: A language and environment for statistical computing*. Foundation for Statistical Computing.

Timarová, S. (2008). Working memory and simultaneous interpreting. Translation and its others: selected papers of the CETRA research seminar in translation studies, Belgium.

Tommola, J., **Laine, M.**, **Sunnari, M.**, & **Rinne, J. O.** (2000). Images of shadowing and interpreting. *Interpreting*, **5**(2), 147–167.

Tzou, Y. Z., **Eslami, Z. R.**, **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**(2), 213–227. https://doi.org/10.1177/1367006911403197.

Ünlü, E. A., & **Şimşek, Ç. S.** (2018). Testing the impact of formal interpreting training on working memory capacity: Evidence from Turkish–English students–interpreters. *Lingua*, **209**, 78–88. https://doi.org/10.1016/j.lingua.2018.04.003.

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.

Wessel, I., **Huntjens, R. J. C.**, & **Verwoerd, J. R. L.** (2010). Cognitive control and suppression of memories of an emotional film. *Journal of Behavior Therapy and Experimental Psychiatry*, **41**(2), 83–89. https://doi.org/10.1016/j.jbtep.2009.10.005.

Zheng, B., **Báez, S.**, **Su, L.**, **Xiang, X.**, **Weis, S.**, **Ibáñez, A.**, & **García, A. M.** (2020). Semantic and attentional networks in bilingual processing: fMRI connectivity signatures of translation directionality. *Brain and Cognition*, **143**, 105584. https://doi.org/10.1016/j.bandc.2020.105584.

https://doi.org/10.1017/S1366728925100898 Published online by Cambridge University Press