NeuroImage 209 (2020) 116519

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Task-specific signatures in the expert brain: Differential correlates of translation and reading in professional interpreters

Martin Dottori [a,b], Eugenia Hesse [a,b,c], Micaela Santilli [a], Martina G. Vilas [a,b], Miguel Martorell Caro [a], Daniel Fraiman [b,c], Lucas Sedeño [a,b], Agustín Ibáñez [a,b,d,e,f], Adolfo M. García [a,b,g,*]

[a] Laboratory of Experimental Psychology and Neuroscience (LPEN), Institute of Cognitive and Translational Neuroscience (INCYT), INECO Foundation, Favaloro University, Buenos Aires, Argentina
[b] National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
[c] Departamento de Matemática, Universidad de San Andrés, Buenos Aires, Argentina
[d] Universidad Autónoma del Caribe, Barranquilla, Colombia
[e] Center for Social and Cognitive Neuroscience (CSCN), School of Psychology, Universidad Adolfo Ibáñez, Santiago de Chile, Chile
[f] Centre of Excellence in Cognition and its Disorders, Australian Research Council (ACR), Sydney, Australia
[g] Faculty of Education, National University of Cuyo (UNCuyo), Mendoza, Argentina

ARTICLE INFO

Keywords:
Bilingualism
Simultaneous interpreting
Expertise
Lexical processing
Electroencephalography
Oscillatory activity

ABSTRACT

Insights on the neurocognitive particularities of expert individuals have benefited from language studies on professional simultaneous interpreters (PSIs). Accruing research indicates that behavioral advantages in this population are restricted to those skills that are directly taxed during professional practice (e.g., translation as opposed to reading), but little is known about the neural signatures of such selective effects. To illuminate the issue, we recruited 17 PSIs and 15 non-interpreter bilinguals and compared behavioral and electrophysiological markers of word reading and translation from and into their native and non-native languages (L1 and L2, respectively). PSIs exhibited greater delta-theta (1–8 Hz) power across all tasks over varying topographies, but these were accompanied by faster performance only in the case of translation conditions. Moreover, neural differences in PSIs were most marked for L2-L1 translation (the dominant interpreting direction in their market), which exhibited maximally widespread modulations that selectively correlated with behavioral outcomes. Taken together, our results suggest that interpreting experience involves distinct neural signatures across reading and translation mechanisms, but that these are systematically related with processing efficiency only in domains that face elevated demands during everyday practice (i.e., L2-L1 translation). These findings can inform models of simultaneous interpreting, in particular, and expert cognitive processing, in general.

1. Introduction

Neurocognitive research on expertise has been fueled by insights from professional simultaneous interpreters (PSIs). These multilinguals are trained to accurately reformulate oral messages from and into their native and foreign languages (L1s and L2s) under strict time constraints (Chernov, 2004; Christoffels and de Groot, 2005; García, 2014), typically with no previous rehearsal. Relative to non-interpreter bilinguals (NIBs), PSIs are characterized by diverse neuroplastic effects (Elmer et al., 2014a; Elmer, Meyer, & Jancke, 2010) and enhanced performance in specific cognitive functions (Becker et al., 2016; Strobach et al., 2015). Within the verbal domain, these behavioral advantages seem restricted to skills that are directly taxed during simultaneous interpreting (SI) –e.g., word translation as opposed to word reading–, suggesting that SI-related changes do not generalize to linguistic processing at large (García et al., 2019). However, there is very little evidence on the neural correlates of these selective linguistic effects, and no study has examined potential associations between such signatures and outward performance. Thus, a major gap exists towards the formulation of multidimensional neurocognitive models in the field. To foster progress in this direction, we conducted the first assessment of oscillatory modulations during translation and reading tasks in PSIs and NIBs.

* Corresponding author. Institute of Cognitive and Translational Neuroscience & CONICET; Pacheco de Melo 1860, C1126AAB, Buenos Aires, Argentina.
E-mail address: adolfomartingarcia@gmail.com (A.M. García).

https://doi.org/10.1016/j.neuroimage.2020.116519
Received 31 July 2019; Received in revised form 23 December 2019; Accepted 2 January 2020
Available online 7 January 2020

Bilingual verbal skills face elevated demands during professional SI. In conference settings, incoming discourse typically exceeds the ideal rate of 95–120 words per minute (Chernov, 2004; Gerver, 1975), with ear-voice spans ranging between 2 and 4 s (Anderson, 1994; Gerver, 1976) and periods of overlapping input and output amounting to 70% of individual sessions (Chernov, 1994). As recently proposed in an integrative review (García et al., 2019), these outstanding linguistic demands can lead to behavioral verbal advantages in PSIs, but only for skills that prove critical in their trade. For example, in the lexical domain, PSIs outperform NIBs in word translation (Christoffels et al., 2006) and language detection (Aparicio et al., 2017), two domains that are fundamental for successful professional performance. Conversely, they have no advantages in lexical decision (Hiltunen et al., 2016) or word reading (Santilli et al., 2018), both skills playing no distinct role in SI relative to other multilingual activities (García et al., 2019). This suggests that SI experience might impact on specific verbal mechanisms in a differential manner.

Nevertheless, relevant neural effects in PSIs seem widespread across language systems at large. These specialists, as well as SI trainees, exhibit structural differences in various brain areas, including perisylvian, frontostriatal, and parietal regions implicated in bilingual cognitive control and linguistic processing (Becker et al., 2016; Elmer et al., 2014a; Hervais-Adelman et al., 2017; Van de Putte et al., 2018). Also, they present increased neurophysiological modulations underlying diverse word-level operations (Elmer and Kuhnis, 2016; Elmer et al., 2010) –e.g., greater theta-band (4–7 Hz) connectivity between the left auditory-related cortex and Broca's area during semantic decision in single- and dual-language conditions (Elmer and Kuhnis, 2016). *Prima facie*, this would indicate that the domain-selectivity of behavioral advantages may not be mirrored by highly circumscribed effects at a cerebral level.

Notably, however, at least some neural patterns in PSIs seem differentially related to specifically taxed domains. Indeed, distinct neurofunctional patterns in this population have been detected for L2-L1 processes (namely, the ones corresponding to the dominant direction in professional practice) relative to L1-L2 and single-language processes (Christoffels et al., 2013; Elmer et al., 2010). In fact, this pattern has also been tracked longitudinally over the course of SI training (Hervais-Adelman et al., 2015a). Therefore, it appears that broad-gauge neurobiological changes along language systems may also be accompanied by distinct patterns for the mechanisms that are more markedly recruited in daily professional practice.

Yet, this notion remains empirically underspecified, as few studies have jointly assessed behavioral and neurofunctional signatures of interpreting experience across intra- and cross-linguistic tasks (Elmer and Kuhnis, 2016; Elmer et al., 2010; Hervais-Adelman et al., 2015a,b; Van de Putte et al., 2018), and none has tested for direct associations between such measures. Moreover, no report seems to have capitalized on the high sensitivity of frequency analysis to address the issue. Unlike other neuroscientific measures, frequency analyses allow disentangling distinct but parallel neural processes which occur simultaneously as an overall task unfolds (Roach and Mathalon, 2008). Crucially, this allows decomposing neurophysiological signals into co-occurring but discernible frequency bands, each of which can index cognitive processes that often prove untraceable without this level of granularity (Kielar et al., 2014). Indeed, by targeting specific bands, studies on L1 (Braunstein et al., 2012) and L2 (Vilas et al., 2019) processing have revealed significant differences (between conditions or groups) that escape other analyses of even the same electrophysiological signals, such as those based on event-related potentials (ERPs). Moreover, frequency analyses can capture subtle effects during word reading (Klimesch et al., 1997; Rohm et al., 2001) and translation (Grabner et al., 2007) processes, and they are particularly sensitive to expertise-related effects across various populations (Behroozmand et al., 2015; Doppelmayr et al., 2008; Pallesen et al., 2015). Hence, frequency analyses afford a particularly promising framework to examine the conjecture raised above.

In this sense, the 1–8 Hz frequency range, which encompasses the delta (0.5–4 Hz) and theta (4–8 Hz) bands, emerges as a sensitive candidate for investigation. Indeed, lexico-semantic processing seems to be indexed by oscillatory changes that spread across both bands (Davidson and Indefrey, 2007; Hald et al., 2006; Kielar et al., 2014) or within either of them (Allefeld et al., 2005; Bastiaansen et al., 2005; Grabner et al., 2007; Molinaro and Lizarazu, 2018; Vilas et al., 2019) in a host of native-language, foreign-language, and even translation-specific processes. Moreover, power increases over different portions of this range have been reported among correlates of expertise in other domains (Behroozmand et al., 2015; Doppelmayr et al., 2008; Pallesen et al., 2015). Accordingly, increased event-related power synchronization across the delta-theta (1–8 Hz) band represents a likely candidate to index expertise effects in PSIs relative to NIBs across linguistic tasks.

Against this background, we conducted the first EEG-based comparison of word reading and translation processes in PSIs and NIBs. Based on previous findings, we hypothesized that selective advantages in translation, indexed by faster response times (RTs) for PSIs, would be accompanied by increased event-related power synchronization across tasks, presumably over the delta-theta band. In addition, we predicted that such differences would be more marked for specifically trained skills (i.e., translation abilities) and selectively related to processing speed (RTs) in them. Moreover, we examined other frequency bands –alpha (8–13 Hz), low beta (13–21 Hz), and high beta (21–34 Hz)– to explore whether potential delta-theta effects are specific to that frequency range or manifest in others as well. Finally, to further test the specificity of this effect, we conducted exploratory analyses of relevant ERPs. Briefly, with this approach, we aimed to better understand the scope of neurocognitive effects in a model of expert language processing.

2. Methods

2.1. Participants

Thirty-four subjects were recruited for the study, but two of them were removed due to excessive noise in the signals. The final sample thus comprised 32 subjects, who partially overlap with those from a previous report (Santilli et al., 2018), namely: 17 PSIs and 15 NIBs. Importantly, this sample size was similar to or larger than those reported in previous EEG studies assessing translation processes and/or differences between PSIs and NIBs (Christoffels et al., 2013; Elmer and Kuhnis, 2016; Elmer et al., 2010; Grabner et al., 2007; Jost et al., 2018; Klein et al., 2018). Moreover, to determine the statistical power of our sample size, we used G*Power 3.1 (Faul et al., 2007) to calculate the minimum effect size detectable given our experimental design and statistical approach (a mixed effects ANOVA, including 32 subjects, 4 measures, and a resulting total sample size of 128). Having set a power of 0.95 and an alpha level of 0.05, we obtained a η_p^2 of 0.06. This effect size is lower than those reported in a previous EEG study comparing PSIs and NIBs on L1, L2, L1-L2, and L2-L1 tasks (Elmer et al., 2010) –namely, a η_p^2 of 0.84 for a main effect of group and a η_p^2 of 0.64 for a group-by-task interaction. Therefore, our study is capable of detecting differences that are even smaller (i.e., harder to trace) than those obtained in similar studies.

All participants were native speakers of Spanish (L1) with high proficiency in English (L2). They were right-handed and had normal or corrected-to-normal vision. None of them reported a history of neurological or psychiatric disease. The PSIs had a mean of 14.65 years of experience ($SD = 12.09$), mainly in the field of conference interpreting. The NIBs were either English teachers or advanced students at an English teacher program, with no experience in interpreting. Data from a previously reported self-report questionnaire (García et al., 2014) showed that the samples were not significantly different in terms of demographic variables (gender and age) and linguistic factors (competence in and weekly exposure to both languages, age of L2 learning, and years of L2 study), with additional tests revealing non-significant differences in

short-term memory, cognitive flexibility, and overall executive skills (all p-values = n.s.). Crucially, however, PSIs were significantly more competent in both interpreting directions and they spent significantly more time engaging in such activities each week (all p-values $< .001$) –although their professional practice mainly required them to interpret from L2 to L1. For details on these variables, including descriptive statistics, p-values, and effect sizes, see Supplementary material 1 (Table S1).

All subjects signed an informed consent and all experimental protocols were performed in accordance with the Declaration of Helsinki. This study was approved by the Bioethics Committee of the National University of Mar del Plata.

2.2. Materials

The study consisted of two previously reported reading and translation tasks, developed on Python programming language (www.python.org) with the Pygame development library (www.pygame.org). Taken together, the tasks involved 384 nouns, half in each language (García et al., 2014). The stimuli were divided into three blocks of 64 items per language, with each block comprising the same number ($n = 16$) of concrete cognates (e.g., *roca*, *rock*), abstract cognates (e.g., *comedia*, *comedy*), concrete noncognates (e.g., *mesa*, *table*), and abstract non-cognates (e.g., *castigo*, *punishment*).[1,2] The Spanish and English blocks were matched for frequency ranking ($p = .97$) and syllabic length ($p = .99$), and blocks within each language were additionally matched for frequency (Spanish: $p = .95$; English: $p = .98$) –data for these variables were extracted from Davies (2008a, 2008b). Stimuli were pseudorandomly distributed within each block, such that items with similar phonological patterns or semantic proximity were at least separated by two intermediate items. Below we describe how these blocks were allocated across the four experimental tasks, namely: L1 reading (L1R), L2 reading (L2R), backward translation (BT, from L2 to L1), and forward translation (FT, from L1 to L2).

2.3. Tasks and procedure

Within-language processing was tested through two previously reported reading tasks (García et al., 2014; Santilli et al., 2018) based on two of the stimulus blocks described above –one in Spanish, for L1R (64 items), and one in English, for L2R (64 items). None of the items in these blocks was a translation equivalent relative to those of the other language. Note, too, that these two blocks were used exclusively for the reading (as opposed to the translation) tasks. Each trial began with a fixation cross (shown for a random period of 100–300 ms) and continued with the target word (which remained visible for 200 ms). Stimuli were presented in white letters (font: Times New Roman; size: 70 pts), centered on the screen against a black background. In each task, participants were instructed to read each word out loud, as fast and accurately as possible, and to press a key as soon as they began articulating their response. The keystroke served both to record RTs and to cue the following trial. Importantly, note that theuse of key-presses allows circumventing phonetic and phonological confounds affecting oral responses (Rastle et al., 2005; Rastle and Davis, 2002) and that both vocal and manual responses prove equally sensitive to particular linguistic effects (Hutson et al., 2013). Accuracy was judged by two separate examiners on a trial-by-trial basis, and the few cases of disagreement were settled by a third examiner. RTs were recorded by a custom-made Python script (www.python.org). Trials were considered invalid if the subject (i) failed to respond, (ii) committed a false start, (iii) uttered a wrong word, or (iv) translated the stimulus instead of reading it. To avoid order-related biases, the reading tasks (L1R, L2R) were counterbalanced across participants. Taken together, both reading tasks lasted approximately 10 min.

Between-language processes were examined via two previously reported translation tasks (García et al., 2014; Santilli et al., 2018) –one for BT and one or FT. BT performance was assessed with one of the two remaining English blocks (64 items), whereas FT performance was examined via one of the two remaining Spanish blocks (64 items). Importantly, these blocks were used exclusively for the translation (as opposed to the reading) tasks. To avoid priming effects between tasks, the blocks used for BT and FT in each subject were chosen so that the items in one language would not be translation equivalents of those in the other language. Also, half the sample performed the BT task with one English block and the other half did so with the other English block (and the same was true of the use of the Spanish blocks for the FT task). The structure and response modality of each trial was exactly identical to that described for the reading tasks. Trials were rejected if the subject (i) failed to respond, (ii) committed a false start, (iii) read the word instead of translating it, or (iv) provided either a wrong or non-predefined translation.[3] As was the case with the reading tasks, the order of BT and FT tasks was counterbalanced across participants. Taken together, the two translation tasks lasted approximately 20 min. The full counterbalancing scheme of all tasks can be found in the Supplementary material 1 (Table S2).

2.4. Behavioral data analysis

Differences in the number of rejected trials, as well as accuracy and RT outcomes, were analyzed via mixed effects ANOVAs including a between-subjects factor (group: PSIS, NIBs) and two within-subjects factors (task: reading, translation; source language: L1, L2) –with subjects as a random factor. In each analysis, a Gaussian error distribution with an identity link function was assumed for the dependent variable and the significance threshold was set to $p < .05$. Post hoc analyses for significant interactions were made via Tukey's HSD test. For each subject, mean accuracy was calculated as the percentage of correct recorded trials per condition, while mean RT was calculated considering only correct responses with latencies below 2000 ms, as in previous works (Fine et al., 2013; Marinus and de Jong, 2011). Effect sizes for main effects and interactions were calculated based on partial eta squared, η_p^2. Depending on the value of this index, effect sizes can be considered small (>0.02), medium (>0.13), or large (>0.26) (Cohen, 1988). In the case of pair-wise comparisons, effect sizes were calculated through Cohen's *d*, an index that also discriminates among small (0–0.20), medium (0.50–0.80), and large (>0.80) effects (Cohen, 1988). All statistical analyses were performed on Statistica 10 (http://www.statsoft.com/). Of

[1] Cognates are words from a given language that have major orthographic and/or phonological overlap with a viable translation equivalent in another language (e.g., *cámara* in Spanish and *camera* in English). Noncognates are words which lack such sublexical overlap relative to their translation equivalents.

[2] Note that our design is not aimed at targeting concreteness or cognate status as factors for analysis. Rather, the inclusion of exemplars representing all possible combinations of these variables sought to ensure that the results could be generalized to lexico-semantic processing at large, especially considering that both concreteness (Barber et al., 2013; Jessen et al., 2000; van Hell and de Groot, 1998a, 1998b) and cognate status (Broersma et al., 2016; Christoffels et al., 2003; Midgley et al., 2011) can modulate behavioral and neural responses during bilingual lexical processing.

[3] This criterion, present in previous reports of the task (García et al., 2014; Santilli et al., 2018), does not imply that one possible translation of a given stimulus is more correct or desirable than another. Rather, it represents a methodological constraint to ensure that the cognate and non-cognate trials that entered the analyses actually reflected processing of such categories. For example, the word *fury* could be reasonably translated as *furia*, *ira*, *rabia* or *enojo*. However, since *fury* has been tagged as an abstract cognate in our stimulus blocks, only its rendition as *furia* (i.e., an abstract cognate) was empirically relevant to the present study.

note, whereas the behavioral data comes from a sample that partially overlaps the one reported in a previous paper (Santilli et al., 2018), its combination with the newly analyzed EEG data allows addressing these study's novel aims, namely: (i) detecting real-time neural differences between groups and (ii) tracking associations between discriminatory neural patterns and performance in specific tasks.

2.5. EEG methods

2.5.1. Acquisition and preprocessing

EEG activity was recorded online during all four tasks for each participant. Signals were acquired through a Biosemi Active-two 128-channel system with pre-amplified sensors and a DC coupling Amplifier. All signals were originally sampled at 1024 Hz, later resampled to 512 Hz, and referenced to the average of all channels. Similarly to previous works (Christoffels et al., 2013; Kielar et al., 2014; Vilas et al., 2019), EEG data were filtered between 0.5 and 45 Hz, and epochs were selected from continuous data in a window from −0.5 s to 1 s around the time of stimulus onset. In line with reported procedures (Vilas et al., 2019), eye movements and blink artifacts were corrected with independent component analysis, remaining artifacts were rejected through visual inspection, and noisy channels were corrected by interpolation. Epochs corresponding to incorrect, invalid, or excessively long (>2000 ms) RTs were excluded from analysis. Considering these criteria, as stated in the "Participants" section, data from two NIBs were excluded from analysis, leading to the final sample of 17 PSIs and 15 NIBs. All EEG signal processing steps were implemented on MATLAB software (vR2016a) through the EEGLAB (v14.1.2) toolbox.

2.5.2. Frequency analysis

Frequency analysis was implemented through EEGLAB software using the Fast Fourier Transform algorithm with a hanning taper. The mean of event-related power synchronization between trials was calculated for each condition and subject, with a baseline of 300 ms before stimulus onset (−300 ms, 0 ms), as in previous works (Hald et al., 2006; Kielar et al., 2014; Vilas et al., 2019). We calculated the power for each trial relative to its respective baseline and then averaged the ensuing values across trials for each subject.

In line with previous works on word translation (Grabner et al., 2007) and bilingual reading (Kielar et al., 2014), power was averaged in different frequency bands of interest. Here we focused on the delta-theta (1–8 Hz), alpha (8–13 Hz), low beta (13–21 Hz), and high beta (21–34 Hz) bands. Also, to gain statistical power and obtain temporally specific results, power changes in each band were calculated for an initial window (0–300 ms) associated with early language access processes, and a later window (300–600 ms) related to various lexico-semantic processes (Grabner et al., 2007; Hald et al., 2006; Vilas et al., 2019; Willems et al., 2008).

2.5.3. Spatial cluster analysis of power across bands

Power differences between groups were assessed in each task through cluster-based topographic analyses considering each frequency band across both time windows. Following the original description by Maris and Oostenveld, 2007 and subsequent neurolinguistic studies (Davidson and Indefrey, 2007; Kielar et al., 2014; Vilas et al., 2019), we implemented this approach through a permutation test. First, for each between-group comparison in each time window and frequency band, we performed a Wilcoxon test (Wilcoxon, 1946) –a univariate non-parametric test that does not assume normal distributions (Sheskin, 2003)– on the power values associated to each electrode, and we obtained the *p*-values corresponding to each electrode. Second, to increase statistical stringency and topographic precision even beyond pioneering reports of this method (Maris and Oostenveld, 2007), we set a threshold of $p \leq .01$ to define clusters of neighboring electrodes with potential differences between groups. With the purpose of identifying topographically consistent differences, clusters were considered significant only if they encompassed more than five electrodes. This approach is empirically advantageous as it can evaluate between-group differences without relying on a priori topographical hypotheses based on ROIs or particular sets of electrodes (Maris and Oostenveld, 2007). As in previous works (Davidson and Indefrey, 2007; Kielar et al., 2014; Maris and Oostenveld, 2007), permutations (5000) were implemented to generate a sample –in our case, for the largest cluster size (cluster-level statistic). This sample was used to determine the significance (Montecarlo *p*-values) of the clusters obtained in the original data by evaluating the proportion of maximum sizes for permutation clusters that are larger than those of our comparison (Maris and Oostenveld, 2007). These clusters were considered significant at a $p < .025$ (assuming an alpha level of .05) relative to the calculated sample, as in previous works (Maris and Oostenveld, 2007; Vilas et al., 2019). Between-group comparisons for each task were corrected for multiple comparisons among the four frequency bands and the two time windows via the false discovery rate (FDR) method, considering up to the three largest clusters discriminating between PSIs and NIBs.

2.6. Correlation between frequency and RT results

In order to establish direct links between significant neural and behavioral patterns, we evaluated whether band-specific power differences were related to outcomes in each task yielding differential performance between groups. To this end, as in previous studies linking EEG modulations with behavioral performance (Melloni et al., 2016), we used Spearman's correlations (a test recommended for non-normal data distributions) to calculate nonlinear associations between the mean power values of the significant clusters and the mean RT of the corresponding condition. This procedure was implemented for all subjects of each group. Correlations were deemed significant if, after an FDR correction (Benjamini and Hochberg, 1995) considering comparisons between each task and group, they yielded a $p < .05$.

2.7. Complementary ERP analysis

As a complementary exploratory analysis, we examined whether between-group differences in each task were also present in modulations of the N400, an ERP shown to track differences between PSIs and NIBs in other linguistic paradigms (Elmer et al., 2010). To this end, for each task separately (L1R, L2R, BT, FT), we averaged the EEG signal over all trials for each electrode and subject (after removing the mean baseline of 300 ms pre-stimulus). This yielded one activity value for each electrode and subject. Then, we used univariate non-parametric Wilcoxon tests to compare the values for each electrode between groups. Using a threshold of $p = .01$, we aimed to determine differential clusters comparing their sizes with those obtained through a permutation ($n = 1000$) distribution, as done for our frequency analyses (for details, see sections 2.5.2 and 2.5.3).

3. Results

3.1. Behavioral results

After rejection of trials with recording errors, the number of remaining trials for accuracy analysis did not differ significantly between groups [$F(90,30) = 0.879, p = .356, \eta_p^2 = 0.01$], tasks [$F(90,30) = 1.004, p = .319, \eta_p^2 = 0.01$], or source languages [$F(90,30) = 1.004$, $p = .319$, $\eta_p^2 = 0.01$]. Moreover, this analysis revealed null interactions between group and task [$F(90,30) = 1.004$, $p = .319$, $\eta_p^2 = 0.01$] and between group and source language [$F(90,30) = 1.138$, $p = .289$, $\eta_p^2 = 0.01$], confirming that the number of remaining trials was similar for all comparisons. For descriptive statistics, see Supplementary material 2 (Table S3).

Accuracy outcomes revealed a significant effect of task [$F(90,30) = 247.69$, $p < .001$, $\eta_p^2 = 0.66$]), with better performance for reading (98.5%) than translation (86.5%). No other significant effects were found

in terms of accuracy (all *p*-values = n.s.). For descriptive statistics, see Supplementary material 2 (Table S4).

As regards RTs, the number of remaining trials upon rejection of latencies greater than 2000 ms did not differ significantly between groups [$F(90,30) = 0.645$, $p = .4284$, $\eta_p^2 = 0.01$] or source languages [$F(90,30) = 0.162$, $p = .6879$, $\eta_p^2 = 0.00$], but it did differ between tasks [$F(90,30) = 237.319$, $p < .001$, $\eta_p^2 = 0.72$], with more remaining trials for reading (96.9%) than translation (84.8%). Moreover, this analysis revealed null interactions between group and task [$F(90,30) = 0.140$, $p = .71$, $\eta_p^2 = 0.01$] and between group and source language [$F(90,30) = 0.138$, $p = .71$, $\eta_p^2 = 0.01$], indicating that the number of trials for RT analysis was similar for all conditions. For descriptive statistics, see Supplementary material 2 (Table S5).

RT outcomes (Figure 1, panel A) revealed a main effect of task [$F(90,30) = 387.32$, $p < .001$, $\eta_p^2 = 0.81$], revealing faster performance for reading ($M = 370.5$ ms) than translation ($M = 608.3$ ms). This effect was qualified by an interaction between group and task [$F(90,30) = 11.0113$, $p < .001$, $\eta_p^2 = 0.11$], with a post hoc analysis (Tukey's HSD test, MSE = 41.6, $df = 30$) revealing faster performance for PSIs than NIBs in translation ($p = .0133$) but not in reading ($p = .489$) tasks. Also significant were the interactions between source language and task [$F(90,30) = 13.7187$, $p < .001$, $\eta_p^2 = 0.13$]. A post hoc analysis (Tukey's HSD test, MSE = 17.1, $df = 30$) showed faster performance for reading than translation when based on L1 stimuli ($p < .001$) and also when based on L2 stimuli ($p < .001$). For descriptive statistics, see Supplementary material 2 (Table S6).

3.2. Frequency results

Given that frequency analyses are performed as between-group comparisons for each task separately, the number of trials rejected due to faulty signals was compared in each case via unpaired two-tailed *t*-

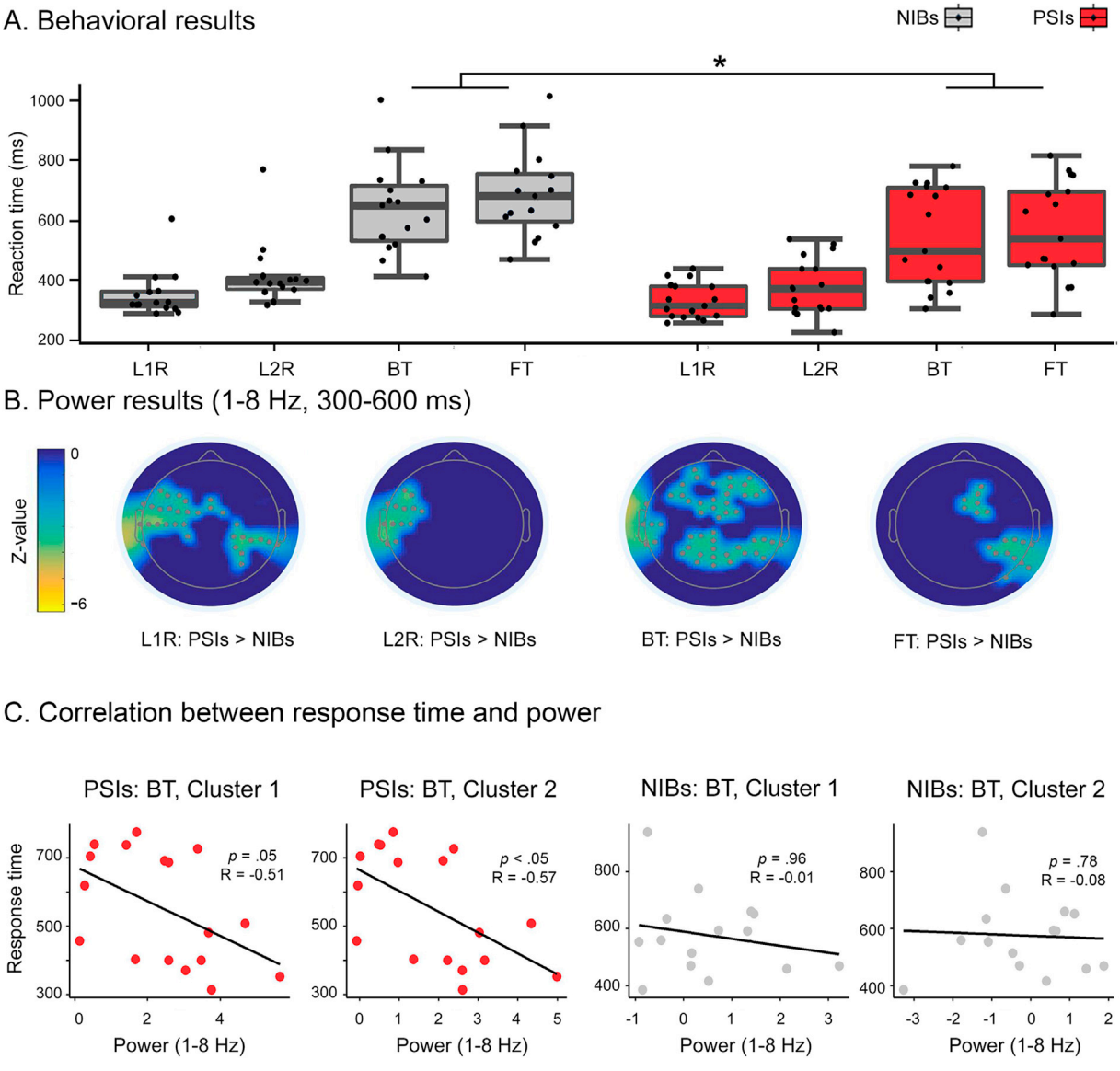

Fig. 1. Response times, power differences, and behavioral-neural associations for PSIs and NIBs during word reading and translation. A. Response times for each task in PSIs and NIBs. The asterisk (*) indicates significant differences. **B.** Clusters yielding significant between-group differences in each task. Results revealed selective and consistent power increases for PSIs over NIBs in the delta-theta band (1–8 Hz) across tasks. The color bar indicates statistics associated to univariate tests of significant electrode clusters. **C.** Scatterplot of the associations between the mean power of significant clusters and response time for BT. Significant associations were found for PSIs (first and second insets), there being no significant associations for NIBs in their respective task-specific clusters (third and fourth insets). BT: backward translation; FT: forward translation; L1R: native-language reading; L2R: foreign-language reading; NIBs: non-interpreter bilinguals; PSIs: professional simultaneous interpreters.

tests. The trials thus rejected was similar between PSIs and NIBs in each task –for descriptive statistics, see Supplementary material 2.1 (Table S7).

Frequency analyses for the reading tasks showed significant clusters discriminating between groups (Figure 1, panel B), characterized by higher delta-theta band (1–8 Hz) power for the PSIs in the later time-window (300–600 ms). For L1R, we observed two significant clusters, one comprising principally frontal and posterior electrodes over the left hemisphere (FDR-corrected $p = .01$), and another one mainly comprised of right posterior electrodes (FDR-corrected $p = .01$). For L2R, the only significant cluster was found over left frontal and posterior electrodes (FDR-corrected $p = .01$), mainly overlapping with the first L1R cluster. None of the remaining combinations of time window and frequency band yielded significant differences at the established thresholds.

As regards translation tasks, given that RTs were greater in NIBs than in PSIs, we first ensured that potential neural activity differences between groups would not be driven by the later responses of NIBs –indeed, longer latencies could involve neural differences across frequency bands due to motoric artifacts rather than condition-specific modulations. To this end, for each task separately, we omitted the top 5% of trials yielding the longest RTs in each NIB and the top 5% of trials yielding the shortest RTs in each PSI. This provided an adequate empirical framework to circumvent the potential biases mentioned above, given that the remaining number of trials in each task and condition was similar for both groups (see Supplementary material 2.1), and their mean RTs did not yield significant between-group differences for either BT [$t(30) = 0.98$, $p = .34$, $d = 0.35$] or FT [$t(30) = 1.39$, $p = .19$, $d = 0.49$].

As observed for reading, frequency results for the translation tasks showed that, relative to NIBs, PSIs exhibited consistently higher power in the delta-theta (1–8 Hz) band across the later time-window (300–600 ms). However, differences were markedly more widespread for BT, with three clusters discriminating between groups: one was distributed across left frontal and posterior sites (FDR-corrected $p = .002$), another one extended principally over bilateral frontal channels (FDR-corrected $p = .002$), and the third was mainly comprised of right posterior electrodes (FDR-corrected $p = .001$). By contrast, FT yielded less distributed between-group differences over two clusters: one consisting of right frontal electrodes (FDR-corrected $p = .02$) and another one spreading over right posterior electrodes (FDR-corrected $p = .02$). None of the remaining combinations of time window and frequency band yielded significant differences at the established thresholds. Notably, these results remained nearly identical when the analyses were rerun without omitting the top 5% of trials yielding the longest RTs in NIBs and the shortest RTs in PSIs –for details, see Supplementary section 2.2 and Figure S1.

Additionally, we explored electrophysiological differences between reading and translation tasks. To this end, we followed the same analysis criteria detailed in sections 2.5.2 and 2.5.3 of the main manuscript (the only difference being that, as required by within-group contrasts, we employed two-tailed unpaired *t*-tests instead of Wilcoxon tests). We focused on the frequency band and time window yielding significant effects in our main analysis (i.e., the delta-theta band in the 300–600 ms window). As in previous neuroscientific studies comparing translation and single-language conditions (Klein et al., 1995; Rinne et al., 2000), we contrasted L1R with FT and L2R with BT –i.e., the conditions that used the same language for their stimuli. Results revealed widespread power increases for the reading over the translation conditions in both groups –see Supplementary material 2.3, Figure S2.

3.3. Exploratory examination of word-category effects

For strictly exploratory purposes, we examined whether between-group differences were driven by specific lexical categories, namely: abstract cognates, abstract noncognates, concrete cognates, concrete noncognates. Analysis of accuracy, RT, and frequency modulations convergently showed that the between-group differences reported above were not driven by any of these four word categories in particular (see Supplementary material 2.4).

3.4. Correlations between RTs and significant power patterns

Given that both translation tasks yielded significant behavioral differences, we explored possible associations between RTs in each of them and the mean power of significant clusters via Spearman correlations –including FDR correction (Figure 1, panel C). For PSIs, we found that RTs for BT had significant negative associations with the frontal ($rho = -0.51$, cluster-corrected $p = .05$) and the right posterior ($rho = -0.57$, cluster-corrected $p < .05$) clusters. No significant association was found between RTs for BT and the remaining cluster ($rho = 0.15$, cluster-corrected $p = .56$) –for details, see Supplementary material 2.5 (Table S10). Neither did this group exhibit any significant associations between FT and underlying neural activity. Finally, results for NIBs revealed no significant associations between RTs and neural activity in either BT (cluster 1: $rho = -0.01$, cluster-corrected $p = .96$; cluster 2: $rho = -0.08$, cluster-corrected $p = .78$; cluster 3: $rho = 0.26$, cluster-corrected $p = .35$) or FT (cluster 1: $rho = -0.03$, cluster-corrected $p = .92$; cluster 2: $rho = 0.06$, cluster-corrected $p = .83$) –for details, see Supplementary material 2.5 (Table S11).

This overall pattern of results remained the same upon rerunning the analyses without omitting the top 5% of trials yielding the longest RTs in NIBs and the shortest RTs in PSIs: following FDR correction, the only significant correlations were those involving BT in PSIs, both in cluster 1 ($rho = -0.5319$, cluster-corrected $p = .0451$) and cluster 2 ($rho = -0.5613$, cluster-corrected $p = .0451$). Every other correlation in PSIs and NIBs remained non-significant after FDR correction –for details, see Supplementary material 2.6 (Table S12 and Table S13).

3.5. Correlations between frequency patterns and years of interpreting experience

Finally, we tested for possible associations between frequency patterns and interpreting experience in the PSIs group. Power values for clusters differentiating PSIs from NIBs in each task were averaged and tested for correlations with years of interpreting experience. We found no significant associations between these metrics in either of our two analytical approaches, that is: neither when removing the 5% of fastest and slowest trials to remove potential motor artifacts ($rho < 0.38$, $p > .13$), nor when retaining all valid trials ($rho < 0.38$, $p > .13$).

3.6. Complementary ERP results

Results from the complementary ERP analyses showed no significant clusters in any of the four tasks when focusing on a canonical N400 window of 200–600 ms (Kutas and Federmeier, 2011). Null results were also obtained for each task when considering a shorter (300–500 ms) and a longer (150–650 ms) window. For details about these results, see Supplementary material 2.4 (Table S14).

4. Discussion

Through direct comparisons of PSIs and NIBs, this study examined how neurocognitive correlates of reading and translation are modulated by experience in SI. While behavioral results revealed interpreter advantages only for translation, frequency analyses showed greater delta-theta power for PSIs across both tasks, irrespective of language of presentation. Notably, however, neural differences in PSIs were most marked for BT, which exhibited maximally widespread modulations that selectively correlated with behavioral performance. Below we discuss these findings and their implications for models of SI and expertise-related effects at large.

In line with other studies (Christoffels et al., 2006), PSIs outperformed non-professional bilinguals in translation but not in

single-language production. This pattern supports the recent claim (García et al., 2019) that, *behaviorally*, linguistic advantages in PSIs are confined to domains taxed during professional practice –e.g., discourse comprehension (Yudes et al., 2013) and semantic error detection (Fabbro et al., 1991; Yudes et al., 2013), as opposed to number counting (Signorelli et al., 2011) and word repetition (Hiltunen et al., 2016). Interestingly, this finding supports the "common demands hypothesis," which suggests that transfer between trained skills and performance in fine-grained tasks depends on similarities between the two of them (Oei and Patterson, 2014, 2015).

Notwithstanding, reading tasks did involve *neural* processing differences between groups. In both languages, PSIs showed higher delta-theta band (1–8 Hz) power in the later time-window (300–600 ms), mainly over left frontal and posterior electrodes –with additional recruitment of right posterior sites for L1R. This aligns with previous studies showing greater engagement of similar topographies for PSIs than NIBs in varied linguistic operations, such as semantic decision (Elmer and Kuhnis, 2016). It would thus seem that SI experience can modulate the neural signatures of processes that are behaviorally unaffected –as observed for other language-experience factors, such as age of L2 acquisition (Klein et al., 2018; Vilas et al., 2019). Here, the distributed power increases observed for PSIs during reading could reflect a differential allocation of cognitive resources for diverse language domains, further suggesting that only those subjected to field-specific practice (such as translation) may reach outwardly detectable enhancements.

These insights are further refined by the neural effects observed during translation tasks. Whereas FT (L1-L2 processing) involved power increases for PSIs over right posterior electrodes, BT (L2-L1 processing) was the only condition yielding widespread bilateral power increases across frontal and posterior scalp sites. Suggestively, those topographies correspond to areas implicated in critical processes of SI, including speech comprehension and production (Abutalebi and Green, 2007; Hervais-Adelman and Babcock, 2019), sound-to-articulation mappings (Wise et al., 1999), and domain-general operations mediating language control and attention allocation (Hervais-Adelman and Babcock, 2019; Hervais-Adelman et al., 2015b). Also, given that our results were not driven by any particular lexical subcategory (as detailed in section 3.3), the observed effects seem to reflect general operations cutting across lexico-semantic processing at large.

In line with this distinctive signature of BT, previous studies comparing PSIs with NIBs have reported differential expertise-related patterns for L2-L1 relative to L1-L2 processes (Elmer et al., 2010). Also, neural differences between language directions have been found in studies assessing PSIs only, including fMRI reports on translation (Rinne et al., 2000) and ERP experiments on semantic decision (Proverbio et al., 2004). Considering that, at least in the Western context, professional SI is predominantly performed in backward rather than forward direction (Bros-Brann, 1976; Donovan, 2004; Gile, 2005; Seleskovitch and Lederer, 1989), the distinctive patterns observed for BT may also be driven by field-specific experience. In fact, specific training in backward SI has been claimed to differentially affect the neural signatures of cross-linguistic processes in L1 and L2 (Elmer et al., 2010). Our finding of uniquely widespread power increases for BT relative to all other conditions aligns well with this idea.

Further support for that interpretation comes from correlation results. Although behavioral advantages were detected in both translation tasks, only BT presented a significant association between power increases and RTs –a pattern that was exclusive to PSIs, as opposed to NIBs. In other words, during BT, the greater the synchronization among distributed hubs distinguishing PSIs from NIBs, the faster the experts' performance. To our knowledge, amid the growing neuroscientific literature on SI expertise (García et al., 2019; Hervais-Adelman and Babcock, 2019), this is the first evidence of a link between neural modulations and task-specific performance in PSIs. Suggestively, however, The PSIs' higher power values were not significantly correlated with their years of interpreting experience. Thus, as previously proposed by Elmer et al., (2010: 151–152), this could indicate that neurophysiological modulations related to SI expertise "are more likely related to the specific training during the SI education rather than to the amount of years of experience."

Note, in this sense, that selective links between neural enhancements and behavioral boosts confined to specifically trained behavior have also been observed in populations with expertise in other fields, such as music (Margulis et al., 2009) and cars (Herzmann and Curran, 2011). In the case of PSIs, continual favoring of L2-L1 over L1-L2 interpreting (Bros--Brann, 1976; Donovan, 2004; Gile, 2005; Seleskovitch and Lederer, 1989) could fine-tune the co-dependence between increased neural recruitment and overall processing speed for BT in particular. In this sense, as claimed elsewhere (García et al., 2019), the linguistic particularities of PSIs do not equally manifest across language mechanisms at large.

Interestingly, all these effects were captured in the delta-theta band (1–8 Hz), with null modulations in every other band tested. Note that oscillations within this frequency range have repeatedly proven sensitive to task-internal manipulations in various linguistic operations (Braunstein et al., 2012; Davidson and Indefrey, 2007; Rohm et al., 2001), including word translation (Grabner et al., 2007). More particularly, increased theta synchronization for PSIs over NIBs has been observed during early stages of lexico-semantic processing, particularly between hubs supporting sound-to-meaning and sound-to-articulation processes (Elmer and Kuhnis, 2016). Thus, greater delta-theta power during task-relevant processes could represent a signature of sustained training. In fact, increased modulations in this and other frequency bands have been identified as a hallmark of expertise in other populations characterized by elevated musical (Pallesen et al., 2015) and athletic (Chuang et al., 2013; Doppelmayr et al., 2008) skills. Therefore, although other frequency bands may capture SI-expertise effects in different non-active conditions, such as resting state (Klein et al., 2018), delta-theta band dynamics may afford sensitive markers of task-specific differences in this population. Interestingly, too, the absence of significant effects in the exploratory N400 analyses highlights the distinct sensitivity of oscillatory measures to capture neurocognitive effects within and between bilingual samples, as shown by previous studies (Vilas et al., 2019).

More particularly, considering that our contrasts between reading and translation tasks consistently yielded widespread power increases for the former in both groups, greater delta-theta power could constitute a marker of less effortful processing. Indeed, reading tasks proved to be behaviorally easier than their translation counterparts, and the more expert group (PSIs) also showed power increases in the same band. Still, this conjecture would need to be tested in specific experiments.

Taken together, these findings offer new insights and constraints for neurocognitive models of SI expertise. First, in line with recent reports (Hervais-Adelman et al., 2017; Klein et al., 2018; Van de Putte et al., 2018), we observed neurobiological patterns underlying varied linguistic tasks in PSIs, even when behavioral performance was not enhanced. This reminds us that a null effect on *one dependent variable* (e.g., RT) does not necessarily entail a null effect of *the independent variable* (expertise) itself, delimiting the scope of conclusions derived exclusively from either behavioral or neural data (García et al., 2019).

Second, L2-L1 processing, which proves dominant for the population under study, was typified by distinctively broad patterns of neural hypersynchronization in proportion to behavioral outcomes. This supports and extends the hypothesis that SI-expertise effects are characterized by *demand-based domain specificity* (García et al., 2019): although neural changes underlie diverse verbal functions in PSIs, they prove differentially marked for those domains more consistently taxed in the profession (here, BT). Importantly, the proposed linguistic nature of the observed effects is further supported by the absence of between-group differences in relevant non-verbal domains –namely, short-term memory, cognitive flexibility, and overall executive performance (Supplementary material 1, Table S1). This pattern is not entirely surprising. Although executive tests have often revealed advantages for PSIs relative to NIBs (e.g., Becker

et al., 2016; Strobach et al., 2015), such differences do not hold for all relevant subdomains (García et al., 2019), and even those domain-general functions reported as enhanced in PSIs, including short-term memory (Babcock and Vallesi, 2017; Christoffels et al., 2006; Stavrakaki et al., 2012) and cognitive flexibility (Yudes et al., 2011), often prove similar between both populations (Henrard and Van Daele, 2017; Köpke and Nespoulous, 2006; Signorelli et al., 2011). Indeed, as shown by graphical modeling outcomes, translation speed and executive processing play independent roles in SI performance (Christoffels et al., 2003). Accordingly, as postulated before, the expertise effects captured herein probably reflect differences in lexico-semantic processing per se, as opposed to discrepancies in executive skills between groups. However, this tentative conclusion should be directly tested via comparisons of oscillatory signatures during executive tasks in PSIs and NIBs (focused, for example, on working memory or cognitive flexibility skills).

Third, note that the differential oscillatory pattern underlying BT spanned vast portions of the scalp. Admittedly, interpreting experience may involve more focal effects for BT in other aspects of brain function, such as reduced hemodynamic activity in the caudate nucleus during SI (Hervais-Adelman et al., 2015a,b). Still, our results suggest that task-specific differences should also be acknowledged to rely on *widespread transient plastic effects*. This would speak to a differential recruitment of diverse cognitive resources for specifically trained functions, inviting new research on their nature and interaction.

More generally, our results attest to the complex and situationally-driven interplay of behavioral and electrophysiological signatures of expertise. Specifically, neural effects cutting across an overarching domain (language processing) may be accompanied by enhanced performance in a distinctively trained function (translation), with direct brain-behavior relations emerging only for the sub-function predominantly taxed in daily practice (BT). Tentatively, it would thus seem that *neurocognitive markers of expertise could present a gradient of convergence*, such that they are potentially dissociable for marginally relevant functions but intimately interrelated for specifically taxed skills. This conjecture opens new possibilities to test and refine models of expertise within and beyond the field of SI proper.

5. Limitations and avenues for further research

This work presents some limitations that invite future studies. First, although our sample had sufficient statistical power and proved larger than or similar to that of previous EEG studies assessing translation processes (Christoffels et al., 2013; Grabner et al., 2007; Jost et al., 2018) or comparing PSIs and NIBs (Elmer and Kuhnis, 2016; Elmer et al., 2010; Klein et al., 2018), it would be desirable to replicate this investigation with more subjects. Second, the frequency analysis was performed in predetermined frequency bands and temporal windows. Whereas this reduces data dimensionality and facilitates the interpretation of results, useful refinements could emerge from data-driven approaches. Third, complementary insights could be gained by examining neural correlates of other tasks beyond the scope of the present study, such as picture naming and verbal fluency –for behavioral approximations to these topics, see (Santilli et al., 2018) and (García et al., 2019). Fourth, note that the present protocol made exclusive use of written stimuli. Although this decision allows circumventing confounds associated to superior sound-processing skills in SIs (Elmer et al., 2014b) while favoring comparability with previous neuroscientific translation studies that also employed visual stimuli (Christoffels et al., 2013; Jost et al., 2018; Price et al., 1999; Quaresima et al., 2002), it would be useful for future studies to employ auditory stimuli in tandem with visual ones so as to better ascertain the scope of linguistic advantages in this population. Fifth, as a complement to word-level assessments, discourse-level tasks could be incorporated to meet the imperative of ecological validity. Sixth, note that the PSIs tested here possessed several years of experience and professional practice. Therefore, it could be interesting to replicate this study in a pre/post design with SI trainees, aiming to see how soon these neurocognitive differences emerge after the onset of systematic practice. In this sense, future research in the field could profit from the inclusion of comprehensive assessments of the participants' translation and interpreting skills, ideally through validated instruments (Schaeffer et al., 2019). Finally, our cross-sectional design does not allow establishing whether present findings were *triggered* by SI experience. Indeed, as noted elsewhere, it could well the case that PSIs possessed distinct neurocognitive profiles prior to entering the field (García et al., 2019). Here lies another reason to apply the present protocol in longitudinal or pre/post studies tracking SI trainees before and after training.

6. Conclusion

This is the first study showing a direct and selective link between behavioral and neural signatures of translation performance in PSIs. Our results support the idea that experience-related effects in this population manifest distinctively for specifically taxed abilities, as suggested by recent models of expertise. Further research in this direction can offer even finer insights on how human neurocognition adapts to recurring demands imposed on linguistic and otherwise cognitive systems.

Data availability statement

All demographic, behavioral, and preprocessed EEG data is publicly available on the Open Science Framework (García et al., 2019).

Declaration of competing interest

None to declare.

CRediT authorship contribution statement

Martin Dottori: Methodology, Software, Formal analysis, Data curation, Writing - original draft. **Eugenia Hesse:** Methodology, Software. **Micaela Santilli:** Investigation, Data curation. **Martina G. Vilas:** Investigation, Data curation. **Miguel Martorell Caro:** Investigation, Data curation. **Daniel Fraiman:** Methodology, Formal analysis. **Lucas Sedeño:** Methodology, Formal analysis, Writing - review & editing. **Agustín Ibáñez:** Methodology, Writing - review & editing, Funding acquisition. **Adolfo M. García:** Conceptualization, Validation, Resources, Methodology, Writing - original draft, Writing - review & editing, Supervision, Project administration.

Acknowledgements

This work was supported by grants from CONICET, FONCYT-PICT (2017-1818, 2017-1820), CONICYT/FONDECYT Regular (1170010), FONDAP 15150012, the Global Brain Health Institute (GBHI-UCSF), and the INECO Foundation.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2020.116519.

References

Abutalebi, J., Green, D., 2007. Bilingual language production: the neurocognition of language representation and control. J. Neurolinguistics 20 (3), 242–275.

Allefeld, C., Frisch, S., Schlesewsky, M., 2005. Detection of early cognitive processing by event-related phase synchronization analysis. Neuroreport 16 (1), 13–16.

Anderson, L., 1994. Simultaneous interpretation: contextual and translational aspects. In: Lambert, S., Moser-Mercer, B. (Eds.), Bridging the Gap: Empirical Research in Simultaneous Interpretation. John Benjamins, Amsterdam, pp. 101–120.

Aparicio, X., Heidlmayr, K., Isel, F., 2017. Inhibition efficiency in highly proficient bilinguals and simultaneous interpreters: evidence from language switching and Stroop tasks. J. Psycholinguist. Res. 46 (6), 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? Biling. Lang. Cognit. 20 (2), 403–417.
Barber, H.A., Otten, L.J., Kousta, S.T., Vigliocco, G., 2013. Concreteness in word processing: ERP and behavioral effects in a lexical decision task. Brain Lang. 125 (1), 47–53. https://doi.org/10.1016/j.bandl.2013.01.005.
Bastiaansen, M.C., van der Linden, M., Ter Keurs, M., Dijkstra, T., Hagoort, P., 2005. Theta responses are involved in lexical-semantic retrieval during language processing. J. Cogn. Neurosci. 17 (3), 530–541. https://doi.org/10.1162/0898929053279469.
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.
Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D.A., Larson, C.R., 2015. Functional role of delta and theta band oscillations for auditory feedback processing during vocal pitch motor control. Front. Neurosci. 9 (109) https://doi.org/10.3389/fnins.2015.00109.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57 (1), 289–300.
Braunstein, V., Ischebeck, A., Brunner, C., Grabner, R.H., Stamenov, M., Neuper, C., 2012. Investigating the influence of proficiency on semantic processing in bilinguals: an ERP and ERD/S analysis. Acta Neurobiol. Exp. 72 (4), 421–438.
Broersma, M., Carter, D., Acheson, D.J., 2016. Cognate costs in bilingual speech production: evidence from language switching. Front. Psychol. 7 https://doi.org/10.3389/fpsyg.2016.01461, 1461-1461.
Bros-Brann, E., 1976. Critical comments on H.C. Barik's article "Interpreters talk a lot, among other things. AIIC Bull. 4 (1), 16–18.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, second ed. Erlbaum, Hillsdale, NJ.
Chernov, G.V., 1994. Message redundancy and message anticipation in simultaneous interpretation. In: Lambert, S., Moser-Mercer, B. (Eds.), Bridging the Gap: Empirical Research in Simultaneous Interpretation. John Benjamins, Amsterdam, pp. 139–153.
Chernov, G.V., 2004. Inference and Anticipation in Simultaneous Interpreting: A Probability-Prediction Model. John Benjamins, Amsterdam.
Christoffels, I., de Groot, A., 2005. Simultaneous interpreting: a cognitive perspective. In: Kroll, J., de Groot, A. (Eds.), Handbook of Bilingualism: Psycholinguistic Approaches. Oxford University Press, New York, pp. 454–479.
Christoffels, I.K., De Groot, A.M., Kroll, J.F., 2006. Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. J. Mem. Lang. 54 (3), 324–345.
Christoffels, I.K., De Groot, A.M.B., Waldorp, L.J., 2003. Basic skills in a complex task: a graphical model relating memory and lexical retrieval to simultaneous interpreting. Biling. Lang. Cognit. 6 (3), 201–211.
Christoffels, I.K., Ganushchak, L., Koester, D., 2013. Language conflict in translation: an ERP study of translation production. J. Cogn. Psychol. 25 (5), 646–664.
Chuang, L.Y., Huang, C.J., Hung, T.M., 2013. The differences in frontal midline theta power between successful and unsuccessful basketball free throws of elite basketball players. Int. J. Psychophysiol. 90 (3), 321–328. https://doi.org/10.1016/j.ijpsycho.2013.10.002.
Davidson, D.J., Indefrey, P., 2007. An inverse relation between event-related and time-frequency violation responses in sentence processing. Brain Res. 1158, 81–92. https://doi.org/10.1016/j.brainres.2007.04.082.
Davies, M., 2008a. Corpus del español: 100 millones de palabras, 1200s-1900s. Retrieved October 13, 2013. http://www.corpusdelespanol.org/.
Davies, M., 2008b. The corpus of contemporary American English: 425 million words, 1990-present. Retrieved October 4, 2013. http://corpus.byu.edu/coca/.
Donovan, C., 2004. European masters Project group: teaching simultaneous interpretation into a B language: preliminary findings. Interpreting 6 (2), 205–216. https://doi.org/10.1075/intp.6.2.06don.
Doppelmayr, M., Finkenzeller, T., Sauseng, P., 2008. Frontal midline theta in the pre-shot phase of rifle shooting: differences between experts and novices. Neuropsychologia 46 (5), 1463–1467. https://doi.org/10.1016/j.neuropsychologia.2007.12.026.
Elmer, S., Hanggi, J., Jancke, L., 2014a. Processing demands upon cognitive, linguistic, and articulatory functions promote grey matter plasticity in the adult multilingual brain: insights from simultaneous interpreters. Cortex 54, 179–189. https://doi.org/10.1016/j.cortex.2014.02.014.
Elmer, S., Klein, C., Kuhnis, J., Liem, F., Meyer, M., Jancke, L., 2014b. Music and language expertise influence the categorization of speech and musical sounds: behavioral and electrophysiological measurements. J. Cogn. Neurosci. 26 (10), 2356–2369. https://doi.org/10.1162/jocn_a_00632.
Elmer, S., Kuhnis, J., 2016. Functional connectivity in the left dorsal stream facilitates simultaneous language translation: an EEG study. Front. Hum. Neurosci. 10, 60. https://doi.org/10.3389/fnhum.2016.00060.
Elmer, Stefan, Meyer, M., Jäncke, L., 2010. Simultaneous interpreters as a model for neuronal adaptation in the domain of language processing. Brain Research 1317, 147–156. https://doi.org/10.1016/j.brainres.2009.12.052.
Fabbro, F., Gran, B., Gran, L., 1991. Hemispheric specialization for semantic and syntactic components of language in simultaneous interpreters. Brain Lang. 41 (1), 1–42.
Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39 (2), 175–191.
Fine, A.B., Jaeger, T.F., Farmer, T.A., Qian, T., 2013. Rapid expectation adaptation during syntactic comprehension. PLoS One 8 (10), e77661. https://doi.org/10.1371/journal.pone.0077661.
García, A.M., 2014. The interpreter advantage hypothesis: preliminary data patterns and empirically motivated questions. Transl. Interpreting Stud, 9, 219–238. https://doi.org/10.1075/tis.9.2.04gar.
García, A.M., Ibanez, A., Huepe, D., Houck, A.L., Michon, M., Lezama, C.G., et al., 2014. Word reading and translation in bilinguals: the impact of formal and informal translation expertise. Front. Psychol. 5, 1302. https://doi.org/10.3389/fpsyg.2014.01302.
García, A.M., Muñoz, E., Kogan, B., 2019. Taxing the bilingual mind: effects of simultaneous interpreting experience on verbal and executive mechanisms. Biling. Lang. Cognit. 1–11. https://doi.org/10.1017/S1366728919000063.
Gerver, D., 1975. A psychological approach to simultaneous interpretation. Meta: Int. Transl. J. 20 (2), 119–128.
Gerver, D., 1976. Empirical studies of simultaneous interpretation: a review and a model. In: Briskin, R.W. (Ed.), Translation: Applications and Research. Gardner Press, New York, pp. 165–207.
Gile, D., 2005. Directionality in conference interpreting: a cognitive view. In: Godijns, R., Hindedael, M. (Eds.), Directionality in Interpreting. The 'Retour' or the Native?, Ghent: Communication and Cognition, pp. 9–26.
Grabner, R.H., Brunner, C., Leeb, R., Neuper, C., Pfurtscheller, G., 2007. Event-related EEG theta and alpha band oscillatory responses during language translation. Brain Res. Bull. 72 (1), 57–65. https://doi.org/10.1016/j.brainresbull.2007.01.001.
Hald, L.A., Bastiaansen, M.C., Hagoort, P., 2006. EEG theta and gamma responses to semantic violations in online sentence processing. Brain Lang. 96 (1), 90–105. https://doi.org/10.1016/j.bandl.2005.06.007.
Henrard, S., Van Daele, A., 2017. Different bilingual experiences might modulate executive tasks advantages: comparative analysis between monolinguals, translators, and interpreters. Front. Psychol. 8, 1870. https://doi.org/10.3389/fpsyg.2017.01870.
Hervais-Adelman, A., Babcock, L., 2019. The neurobiology of simultaneous interpreting: where extreme language control and cognitive control intersect. Biling. Lang. Cognit. 1–12. https://doi.org/10.1017/s1366728919000324.
Hervais-Adelman, A., Moser-Mercer, B., Golestani, N., 2015a. Brain functional plasticity associated with the emergence of expertise in extreme language control. Neuroimage 114, 264–274.
Hervais-Adelman, A., Moser-Mercer, B., Michel, C.M., Golestani, N., 2015b. fMRI of simultaneous interpretation reveals the neural basis of extreme language control. Cerebr. Cortex 25 (12), 4727–4739. https://doi.org/10.1093/cercor/bhu158.
Hervais-Adelman, A., Moser-Mercer, B., Murray, M.M., Golestani, N., 2017. Cortical thickness increases after simultaneous interpretation training. Neuropsychologia 98, 212–219. https://doi.org/10.1016/j.neuropsychologia.2017.01.008.
Herzmann, G., Curran, T., 2011. Experts' memory: an ERP study of perceptual expertise effects on encoding and recognition. Mem. Cogn. 39 (3), 412–432. https://doi.org/10.3758/s13421-010-0036-1.
Hiltunen, S., Pääkkönen, R., Vik, G.-V., Krause, C.M., 2016. On interpreters' working memory and executive control. Int. J. Biling. 20 (3), 297–314. https://doi.org/10.1177/1367006914554406.
Hutson, J., Damian, M.F., Spalek, K., 2013. Distractor frequency effects in picture–word interference tasks with vocal and manual responses. Lang. Cogn. Process. 28 (5), 615–632. https://doi.org/10.1080/01690965.2011.605599.
Jessen, F., Heun, R., Erb, M., Granath, D.O., Klose, U., Papassotiropoulos, A., Grodd, W., 2000. The concreteness effect: evidence for dual coding and Context availability. Brain Lang. 74 (1), 103–112. https://doi.org/10.1006/brln.2000.2340.
Jost, L.B., Radman, N., Buetler, K.A., Annoni, J.M., 2018. Behavioral and electrophysiological signatures of word translation processes. Neuropsychologia 109, 245–254. https://doi.org/10.1016/j.neuropsychologia.2017.12.034.
Kielar, A., Meltzer, J.A., Moreno, S., Alain, C., Bialystok, E., 2014. Oscillatory responses to semantic and syntactic violations. J. Cogn. Neurosci. 26 (12), 2840–2862. https://doi.org/10.1162/jocn_a_00670.
Klein, C., Metz, S.I., Elmer, S., Jancke, L., 2018. The interpreter's brain during rest - hyperconnectivity in the frontal lobe. PLoS One 13 (8), e0202600. https://doi.org/10.1371/journal.pone.0202600.
Klein, D., Milner, B., Zatorre, R.J., Meyer, E., Evans, A.C., 1995. The neural substrates underlying word generation: a bilingual functional-imaging study. Proc. Natl. Acad. Sci. U.S.A. 92 (7), 2899–2903.
Klimesch, W., Doppelmayr, M., Pachinger, T., Russegger, H., 1997. Event-related desynchronization in the alpha band and the processing of semantic information. Brain Res. 6 (2), 83–94.
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.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.
Margulis, E.H., Mlsna, L.M., Uppunda, A.K., Parrish, T.B., Wong, P.C., 2009. Selective neurophysiologic responses to music in instrumentalists with different listening biographies. Hum. Brain Mapp. 30 (1), 267–275. https://doi.org/10.1002/hbm.20503.
Marinus, E., de Jong, P.F., 2011. Dyslexic and typical-reading children use vowel digraphs as perceptual units in reading. Q. J. Exp. Psychol. 64 (3), 504–516. https://doi.org/10.1080/17470218.2010.509803.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.
Melloni, M., Billeke, P., Baez, S., Hesse, E., de la Fuente, L., Forno, G., et al., 2016. Your perspective and my benefit: multiple lesion models of self-other integration strategies during social bargaining. Brain 139 (11), 3022–3040. https://doi.org/10.1093/brain/aww231.
Midgley, K.J., Holcomb, P.J., Grainger, J., 2011. Effects of cognate status on word comprehension in second language learners: an ERP investigation. J. Cogn. Neurosci. 23 (7), 1634–1647. https://doi.org/10.1162/jocn.2010.21463.

Molinaro, N., Lizarazu, M., 2018. Delta(but not theta)-band cortical entrainment involves speech-specific processing. Eur. J. Neurosci. 48 (7), 2642–2650. https://doi.org/10.1111/ejn.13811.

Oei, A.C., Patterson, M.D., 2014. Are videogame training gains specific or general? Front. Syst. Neurosci. 8, 54. https://doi.org/10.3389/fnsys.2014.00054.

Oei, A.C., Patterson, M.D., 2015. Enhancing perceptual and attentional skills requires common demands between the action video games and transfer tasks. Front. Psychol. 6, 113. https://doi.org/10.3389/fpsyg.2015.00113.

Pallesen, K.J., Bailey, C.J., Brattico, E., Gjedde, A., Palva, J.M., Palva, S., 2015. Experience drives synchronization: the phase and amplitude dynamics of neural oscillations to musical chords are differentially modulated by musical expertise. PLoS One 10 (8), e0134211. https://doi.org/10.1371/journal.pone.0134211.

Price, C.J., Green, D.W., Von Studnitz, R., 1999. A functional imaging study of translation and language switching. Brain 122 (12), 2221–2235.

Proverbio, A.M., Leoni, G., Zani, A., 2004. Language switching mechanisms in simultaneous interpreters: an ERP study. Neuropsychologia 42 (12), 1636–1656. https://doi.org/10.1016/j.neuropsychologia.2004.04.013.

Quaresima, V., Ferrari, M., van der Sluijs, M.C., Menssen, J., Colier, W.N., 2002. Lateral frontal cortex oxygenation changes during translation and language switching revealed by non-invasive near-infrared multi-point measurements. Brain Res. Bull. 59 (3), 235–243.

Rastle, K., Croot, K.P., Harrington, J.M., Coltheart, M., 2005. Characterizing the motor execution stage of speech production: consonantal effects on delayed naming latency and onset duration. J. Exp. Psychol. Hum. Percept. Perform. 31 (5), 1083–1095. https://doi.org/10.1037/0096-1523.31.5.1083.

Rastle, K., Davis, M.H., 2002. On the complexities of measuring naming. J. Exp. Psychol. Hum. Percept. Perform. 28 (2), 307–314. https://doi.org/10.1037/0096-1523.28.2.307.

Rinne, J.O., Tommola, J., Laine, M., Krause, B.J., Schmidt, D., Kaasinen, V., et al., 2000. The translating brain: cerebral activation patterns during simultaneous interpreting. Neurosci. Lett. 294 (2), 85–88.

Roach, B.J., Mathalon, D.H., 2008. Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. Schizophr. Bull. 34 (5), 907–926. https://doi.org/10.1093/schbul/sbn093.

Rohm, D., Klimesch, W., Haider, H., Doppelmayr, M., 2001. The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. Neurosci. Lett. 310 (2–3), 137–140.

Santilli, M., Gonzalez Vilas, M., Mikulan, E., Martorell Caro, M., Muñoz, E., Sedeño, L., et al., 2018a. Bilingual memory, to the extreme: lexical processing in simultaneous interpreters. Biling. Lang. Cognit. https://doi.org/10.1017/S1366728918000378.

Schaeffer, M., Huepe, D., Hansen-Schirra, S., Hofmann, S., Muñoz, E., Kogan, B., et al., 2019. The Translation and Interpreting Competence Questionnaire: an Online Tool for Research on Translators and Interpreters. Perspectives 1–19. https://doi.org/10.1080/0907676X.2019.1629468.

Seleskovitch, D., Lederer, M., 1989. Pédagogie raisonnée de l'interprétation. Didier Erudition, Paris.

Sheskin, D.J., 2003. Handbook of Parametric and Nonparametric Statistical Procedures. crc Press.

Signorelli, T.M., Haarmann, H.J., Obler, L.K., 2011. Working memory in simultaneous interpreters: effects of task and age. Int. J. Biling. 16 (2), 198–212. https://doi.org/10.1177/1367006911403200.

Stavrakaki, S., Megari, K., Kosmidis, M.H., Apostolidou, M., Takou, E., 2012. Working memory and verbal fluency in simultaneous interpreters. J. Clin. Exp. Neuropsychol. 34 (6), 624–633. https://doi.org/10.1080/13803395.2012.667068.

Strobach, T., Becker, M., Schubert, T., Kuhn, S., 2015. Better dual-task processing in simultaneous interpreters. Front. Psychol. 6, 1590. https://doi.org/10.3389/fpsyg.2015.01590.

Van de Putte, E., De Baene, W., García-Penton, 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 Hell, J.G., de Groot, A.M.B., 1998a. Conceptual representation in bilingual memory: effects of concreteness and cognate status in word association. Biling. Lang. Cognit. 1, 193–211.

van Hell, J., de Groot, A.M.B., 1998b. Disentangling context availability and concreteness in lexical decision and word translation. Q. J. Exp. Psychol. Sect. A 51 (1), 41–63. https://doi.org/10.1080/713755752.

Vilas, M.G., Santilli, M., Mikulan, E., Adolfi, F., Martorell Caro, M., Manes, F., et al., 2019. Reading Shakespearean tropes in a foreign tongue: age of L2 acquisition modulates neural responses to functional shifts. Neuropsychologia 124, 79–86. https://doi.org/10.1016/j.neuropsychologia.2019.01.007.

Wilcoxon, F., 1946. Individual comparisons of grouped data by ranking methods. J. Econ. Entomol. 39, 269.

Willems, R.M., Oostenveld, R., Hagoort, P., 2008. Early decreases in alpha and gamma band power distinguish linguistic from visual information during spoken sentence comprehension. Brain Res. 1219, 78–90. https://doi.org/10.1016/j.brainres.2008.04.065.

Wise, R.J., Greene, J., Buchel, C., Scott, S.K., 1999. Brain regions involved in articulation. Lancet 353 (9158), 1057–1061. https://doi.org/10.1016/s0140-6736(98)07491-1.

Yudes, C., Macizo, P., Bajo, T., 2011. The influence of expertise in simultaneous interpreting on non-verbal executive processes. Front. Psychol. 2.

Yudes, C., Macizo, P., Morales, L., Bajo, M., 2013. Comprehension and error monitoring in simultaneous interpreters. Appl. Psycholinguist. 34, 1039–1057.