Neuroscience

The bilingual lexicon, back and forth: Electrophysiological signatures of translation asymmetry
--Manuscript Draft--

Manuscript Number:	NSC-21-0879R1
Article Type:	Research Paper
Section/Category:	Behavioral and Cognitive Neuroscience
Keywords:	Bilingualism; translation asymmetry; oscillations; functional connectivity; brain-behavior correlations
Corresponding Author:	Adolfo García University of California San Francisco Ciudad Autónoma de Buenos Aires, Buenos Aires ARGENTINA
First Author:	Gonzalo Pérez
Order of Authors:	Gonzalo Pérez
	Eugenia Hesse
	Martín Dottori
	Agustina Birba
	Lucía Amoruso
	Miguel Martorell Caro
	Agustín Ibáñez
	Adolfo García
Abstract:	Mainstream theories of first- and second-language (L1, L2) processing in bilinguals are crucially informed by word translation research. A core finding is the translation asymmetry effect, typified by slower performance during forward translation (FT, from L1 into L2) than backward translation (BT, from L2 into L1). Yet, few studies have explored its neural bases and none has employed (de)synchronization measures, precluding the integration of bilingual memory models with novel neural (de)coupling accounts of word processing. Here, 27 proficient Spanish-English bilinguals engaged in FT and BT of single words as we obtained high-density EEG recordings to perform cluster-based oscillatory and non-linear functional connectivity analyses. Relative to BT, FT yielded slower responses, higher frontal theta (4-7 Hz) power in an early window (0-300 ms), reduced centro-posterior lower-beta (14-20 Hz) and centro-frontal upper-beta (21-30 Hz) power in a later window (300-600 ms), and lower fronto-parietal connectivity below 10 Hz in the early window. Also, the greater the behavioral difference between FT and BT, the greater the power of the early theta cluster for FT over BT. These results reveal key (de)coupling dynamics underlying translation asymmetry, offering frequency-specific constraints for leading models of bilingual lexical processing.
Response to Reviewers:	Manuscript number: NSC-21-0879 Article title: The bilingual lexicon, back and forth: Electrophysiological signatures of translation asymmetry Journal title: Neuroscience Dear Prof Sanes, I hope this message finds you well. I greatly appreciate your diligence in handling our submission. On behalf of my co-authors, I would also like to thank you and the reviewers for making such praiseful and useful comments on our work. Their recommendations have enabled us to produce a more compelling manuscript. Please find below a detailed response to the issues raised. We look forward to the evaluation of this new version.

The bilingual lexicon, back and forth:
Electrophysiological signatures of translation asymmetry

Gonzalo Pérez,[1] Eugenia Hesse,[1,2] Martín Dottori,[1,2] Agustina Birba,[1,2]

Lucía Amoruso,[3,4] Miguel Martorell Caro,[2] Agustín Ibáñez,[1,2,5,6] Adolfo M. García[1,2,5,7,*]

[1] Cognitive Neuroscience Center (CNC), Universidad de San Andrés, Buenos Aires, Argentina

[2] National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

[3] Basque Center on Cognition, Brain and Language (BCBL), San Sebastian, Spain

[4] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

[5] Global Brain Health Institute, University of California San Francisco, San Francisco, California, United States; and Trinity College Dublin, Dublin, Ireland

[6] Latin American Brain Health (BrainLat), Universidad Adolfo Ibáñez, Santiago, Chile

[7] Departamento de Lingüística y Literatura, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago, Chile

* **Corresponding author:** Adolfo M. García, Ph. D., Universidad de San Andrés & CONICET; Vito Dumas 284, B1644BID Victoria, Buenos Aires, Argentina; Phone: (54-11) 4725-7000. E-mail: adolfo. garcia@gbhi.com

Abstract

Mainstream theories of first- and second-language (L1, L2) processing in bilinguals are crucially informed by word translation research. A core finding is the translation asymmetry effect, typified by slower performance during forward translation (FT, from L1 into L2) than backward translation (BT, from L2 into L1). Yet, few studies have explored its neural bases and none has employed (de)synchronization measures, precluding the integration of bilingual

1

memory models with novel neural (de)coupling accounts of word processing. Here, 27 proficient Spanish-English bilinguals engaged in FT and BT of single words as we obtained high-density EEG recordings to perform cluster-based oscillatory and non-linear functional connectivity analyses. Relative to BT, FT yielded slower responses, higher frontal theta (4-7 Hz) power in an early window (0-300 ms), reduced centro-posterior lower-beta (14-20 Hz) and centro-frontal upper-beta (21-30 Hz) power in a later window (300-600 ms), and lower fronto-parietal connectivity below 10 Hz in the early window. Also, the greater the behavioral difference between FT and BT, the greater the power of the early theta cluster for FT over BT. These results reveal key (de)coupling dynamics underlying translation asymmetry, offering frequency-specific constraints for leading models of bilingual lexical processing.

INTRODUCTION

Bilingual memory research has thriven by studying varied interactions between first-language (L1) and second-language (L2) processes (French and Jacquet, 2004; Abutalebi and Green, 2007). An influential finding is the translation asymmetry effect, typified by different cognitive demands for forward translation (FT, from L1 to L2) and backward translation (BT, from L2 to L1) (Kroll et al., 1994; French and Jacquet, 2004; Kroll et al., 2010; Ibrahim et al., 2017). Numerous works have examined this phenomenon through behavioral methods (Kroll et al., 1994; Kroll et al., 2010; Poarch et al., 2015; Ibrahim et al., 2017), prompting classical (Kroll et al., 1994) and recent (Dijkstra et al., 2019) models of bilingual lexical processing. Conversely, few studies have incorporated neuroscientific approaches (García, 2013, 2019) and none has leveraged time-sensitive brain synchrony measures indexing other cognitive distinctions in bilingualism research (Grabner et al., 2007; Elmer and Kühnis, 2016; Vilas et al., 2019; Birba et al., 2020). Moreover, no single experiment has explored direct associations between behavioral and neural signatures of the effect. A divide thus exists between mainstream bilingual memory models and novel brain (de)coupling accounts of word processing. To bridge these gaps, we examined behavioral and neurophysiological (oscillatory and functional connectivity) markers of translation asymmetry and explored potential correlations between them.

2

The ability to translate between languages is a concomitant of bilingualism (Harris and Sherwood, 1978; Malakoff, 1992; García, 2019). Yet, underlying demands vary depending on language directionality. Typically, FT yields longer response times (RTs) than BT (Sáchez-Casas et al., 1992; Degroot et al., 1994; Kroll et al., 1994; De Groot and Poot, 1997; Cheung et al., 1998; Kroll et al., 2010; Poarch et al., 2015; Ibrahim et al., 2017). This translation asymmetry effect has molded the notion of weaker connections from L1 to L2 words than vice versa –a cornerstone of bilingual memory accounts, from the pioneering Revised Hierarchical Model (Kroll et al., 2010) to the contemporary Multilink model (Dijkstra et al., 2019). Limited studies show that FT and BT can be doubly dissociated following brain damage (García, 2013; García, 2015) and that the former elicits greater amplitude in attention-sensitive event-related potentials (Christoffels et al., 2013) as well as enhanced activation (Rinne et al., 2000; Tommola et al., 2000; Quaresima et al., 2002) and functional connectivity (Zheng et al., 2020) along temporal, parietal, and fronto-basal regions subserving lexico-semantic and cognitive control processes. Yet, such evidence fails to reveal whether and how translation asymmetry hinges on *neural (de)synchronization patterns* known to mediate fast-changing linguistic (Grabner et al., 2007; Kielar et al., 2014; Grady et al., 2015; Pérez et al., 2015; Elmer and Kühnis, 2016; García et al., 2016, 2020; Birba et al., 2020; Moguilner et al., 2021) and executive (Grundy et al., 2017a; Grundy et al., 2017b; Litcofsky and Van Hell, 2017) operations. This limits its integration with thriving trends in bilingualism research (Vilas et al., 2019; Birba et al., 2020; Sulpizio et al., 2020; Fan et al., 2021), while adding to the divide between psycholinguistic (Kroll et al., 1994; Dijkstra and Van Heuven, 1998; Dijkstra and Van Heuven, 2002) and neurocognitive (Abutalebi and Green, 2007; Kroll et al., 2013) models in the field.

This gap can be bridged with electroencephalographic (EEG) oscillatory and functional connectivity metrics, which capture ongoing neural dynamics that escape other techniques (Buzsáki and Draguhn, 2004; Buzsaki, 2006; Rubinov and Sporns, 2010, 2011; Friston, 2011; Mišić and Sporns, 2016). While oscillatory measures register transient (de)couplings between cortical cell assemblies (Buzsáki and Draguhn, 2004; Buzsaki, 2006), functional connectivity metrics capture statistical co-dependencies indexing segregated neural processing (Rubinov and Sporns, 2010; Friston, 2011). In particular, modulations of both measures in the theta, alpha, and beta bands seem sensitive to fine-grained effects during L1 and L2 processing (Weiss et al., 2005; Bastiaansen et al., 2010; Bialystok et al.; Kielar et al., 2014; Pérez et al., 2015; Elmer and Kühnis, 2016; Vilas et al., 2019; Birba et al., 2020), as seen even in studies that examined FT and BT separately (Grabner et al., 2007; Dottori et al., 2020). Indeed,

3

modulations in some such bands (viz., theta) correlate positively with word translation speed, suggesting a putative role during the task (Dottori et al., 2020). Thus, key multidimensional insights on translation asymmetry could be gained by examining its oscillatory and functional connectivity signatures in these frequency bands, and their correlation with the effect's behavioral manifestation. Such is the aim of the present study.

We asked proficient Spanish-English bilinguals to perform validated (Santilli et al., 2019; Dottori et al., 2020) FT and BT tasks (alongside L1 and L2 reading tasks for control purposes) as we obtained high-density EEG recordings for cluster-level oscillatory and non-linear functional connectivity analyses. We raised three hypotheses. First, we predicted that FT would elicit longer RTs than BT. Second, we hypothesized that FT and BT would yield differential oscillatory and functional connectivity patterns in the theta, alpha, and/or beta bands. Finally, we anticipated that such differential modulations would correlate with the effect's behavioral manifestation. Briefly, with this approach, we aimed to shed novel light on central phenomenon within bilingual memory research.

EXPERIMENTAL PROCEDURES

Participants

The sample comprised 27 right-handed participants (23 women), yielding adequate power for the proposed analyses (Supplementary material 1). Participants had a mean age of 33.46 (SD = 11.70), normal or corrected-to-normal vision, and no history of neurological or psychiatric disease. All were native speakers of Spanish (L1) with high proficiency in English (L2). Their age of initial L2 exposure ranged from 0 to 13 (mean = 7.88; SD = 3.24). Self-ratings on a scale from 1 (complete inability) to 7 (strong ability) revealed high and similar levels of competence for L1 (6.73, SD = 0.45) and L2 (6.38, SD = 0.57), as well as BT (5.50, SD = 0.95) and FT (5.04; SD = 0.89). All participants signed an informed consent, and all experimental protocols were performed in accordance with the Declaration of Helsinki. This study was approved by the institutional ethics committee.

Experimental protocol

We employed a previously reported task (García et al., 2014; Santilli et al., 2019), involving four conditions: forward translation (FT, from L1 to L2), backward translation (BT, from L2 to L1), L1 reading (L1R), L2 reading (L2R) –Figure 1A. Our focus was on the translation tasks, which can directly capture the asymmetry effect. The reading tasks were included to test for potential differences in more basic single-language processes that are comprised within translation acts (e.g., written word perception and comprehension, spoken word production)

4

but do not presume interlingual reformulation –namely, the distinguishing feature of translation.

The stimulus set comprised 384 nouns, half in each language, grouped into three blocks of 64 items per language. Each block had 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*). 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, 2002, 2010).

Translation tasks were based on two Spanish blocks (for FT) and two English blocks (for BT). To avoid translation priming effects, none of the items in the FT block had an equivalent in the corresponding BT block. Also, half of 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 reading tasks were based on the two remaining blocks (in Spanish for L1R and in English for L2R). None of these items were translation equivalents of each another. To avoid order-related biases, the four tasks were counterbalanced across participants.

In all tasks, each trial began with a fixation cross (visible during a random period of 100 to 300 ms). Then, a target word appeared centered on the screen for about 200 ms in white letters (font: Times New Roman; size: 70 pt), against a black background. In the reading tasks, participants were asked to read out loud the words on the screen as quickly and accurately as possible. In the translation tasks, they had to translate the given words, also as quickly and accurately as possible. Before that, participants were instructed to press a key when they felt ready to give their response. This action served both to record RTs and to cue the following trial. This procedure has been recommended as an alternative to voice-based measures, given the impact of phonological discrepancies across lists on articulation onset (García et al., 2014; Santilli et al., 2019; Dottori et al., 2020).

Overall, the protocol lasted roughly 30 minutes. All tasks were implemented in Python (www.python.org) with the Pygame development library (www.pygame.org). The same computer program was used to record the RTs in every trial.

Behavioral data analysis

Behavioral data was analyzed following reported procedures for the same task (Dottori et al., 2020). First, we calculated mean accuracy, for each subject, as the proportion of correct trials in each condition separately. Trials were considered invalid if the participant: (i) failed to respond, (ii) committed a false start, (iii) uttered a wrong word, (iv) performed the wrong task (reading a stimulus in the translation tasks or vice versa, or, (v) provided either a wrong or non-predefined translation.[1] 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. We then calculated mean RT values, considering only trials with correct answers, and excluding those with latencies above 2000 ms, and more than 3 standard deviations apart from the whole group's mean. The number of rejected trials did not differ significantly between FT and BT or between L2R and L1R (Supplementary material 2, Table S1).

To analyze accuracy and RT data, we implemented two separate 2x2 fixed-effects repeated measures ANOVAs, with task (reading and translation) and stimulus language (L1 and L2) as fixed factors. Interaction effects were analyzed via Tukey post-hoc tests (Abdi and Williams, 2010). Alpha levels were set to .05. Effect sizes were calculated with partial eta squared (η_p^2) for main and interaction effects, and Cohen's *d* for post-hoc pairwise comparisons. These analyses were performed with JASP, v 0.10.2.0 (Love et al., 2019).

EEG acquisition and preprocessing

EEG acquisition and preprocessing steps replicated previous reports with the same tasks (Dottori et al., 2020). EEG recordings were obtained through a Biosemi Active-two 128-channel system with pre-amplified sensors and a DC coupling amplifier. Originally, signals were sampled at 1024 Hz, later downsampled to 512 Hz, and finally referenced to the average of all channels. Following previous studies (Christoffels et al., 2013; Kielar et al., 2014; Vilas et al., 2019), EEG data was band-pass filtered between 0.5 and 45 Hz. Epochs were defined within a time window extending from -0.3 to 1s relative to stimulus onset. Invalid trials were excluded from further analysis. In line with reported procedures, artifacts, such as eye movements and blinks, were corrected using independent component analysis (Vilas et al.,

---

[1] This exclusion criterion, also applied in previous analyses of the same task (Santilli et al., 2019; Dottori et al., 2020), does not suggest that one translation of a particular word is more accurate than another. Rather, it imposes a methodological constraint with the aim to ensure that cognate and non-cognate stimulus are actually processed as such. For example, the word *fury* could be translated as *furia*, *ira*, *rabia* or *enojo*. Nevertheless, *fury* has been tagged as an abstract cognate in our stimulus blocks, and, hence, only its translation as *furia* (i.e., an abstract cognate) was empirically relevant to the present study.

2019; Birba et al., 2020; Dottori et al., 2020). Remaining artifacts were rejected through visual inspection, and noisy channels were interpolated (Birba et al., 2020; García et al., 2020).

Frequency analysis

Spectral analyses were performed via the Fast Fourier Transform with a Hanning taper of 250 ms, using Fieldtrip (Maris and Oostenveld, 2007). Mean trial event-related power synchronization was calculated per subject for each condition. Spectral power within each trial was baseline-normalized and converted to decibels (dB).

We targeted the frequency bands considered in a previous EEG study of word translation: theta (4-7 Hz), alpha (8-13 Hz), lower-beta (14-20 Hz), and upper-beta (21-30 Hz) (Dottori et al., 2020). For each band, as in such study (Dottori et al., 2020), we considered the power elicited in two windows capturing early (0-300 ms) and later (300-600 ms) lexico-semantic processes (Hald et al., 2006; Grabner et al., 2007; Willems et al., 2008; Vilas et al., 2019). For every combination of frequency band and time window, we calculated the mean power for every subject and performed statistical comparisons between FT and BT (to capture power signatures of the translation asymmetry effect) and between L2R and L1R (to rule out potential differences in relevant single-language processes).

Following reported procedures (Vilas et al., 2019; Dottori et al., 2020), we used a cluster-based statistical analysis (Maris and Oostenveld, 2007) to examine differences between experimental conditions in each combination of frequency band and time window. In this approach, signals are averaged in each time window (early and later) and frequency band. This yields one power value per electrode per subject in each experimental condition. Statistical comparisons are performed across participants to identify electrodes that exhibit significant differences between conditions via two-tailed t-tests at $p_{elec} < .05$. Those electrodes are, then, grouped in clusters based on their Euclidean distance. For every cluster observed in the data, a permutation test is performed to generate a histogram of relevant cluster-level statistics (here, the largest sum of the t-values of all electrodes forming the clusters observed in each permutation). The p-value of each cluster is estimated as the proportion of permutations that yielded cluster-level statistics greater than that of the corresponding cluster. Clusters with a $p_{clus} < .05$ are considered significant. For plotting purposes, we generated a binary mask based on the electrodes forming the significant clusters and applied it to the topographic plot corresponding to the power average across participants, for the combination of time window and frequency band associated with the specific cluster. This non-parametric method circumvents the multiple comparisons problem without the need to previously define

7

topographical regions of interest (ROIs) comprising particular sets of electrodes (Maris and Oostenveld, 2007). As in previous works (Dottori et al., 2020), 5000 permutations were implemented to generate the histograms of the cluster-level statistics.

Each cluster's *p*-value was calculated as the proportion of random partitions that generated a cluster-level statistic more extreme than the sum of *t*-values in the observed clusters. Clusters were considered significant if they had a value of p_{clus} <.05. To reject smaller, less significant clusters, as in previous reports of the same task (Dottori et al., 2020), we only considered those comprising more than five electrodes.

Functional connectivity analysis

Functional connectivity was quantified through weighted Symbolic Mutual Information (wSMI) (King et al., 2013), a metric that captures non-linear coupling during lexico-semantic processes in different populations (Hesse et al., 2019; García et al., 2020), including bilinguals (Birba et al., 2020). This method provides a measure of information sharing between two signals over a time interval (King et al., 2013). To this end, the signals are first transformed into a set of discrete symbols, consisting of *k* points, separated by a fixed time interval *τ*. Then, the entropy of each transformed signal is calculated, as well as their joint entropy. These values are then used to obtain the mutual information coefficient for every pair of signals (i.e., signals registered at any pair of electrodes). After that, binary weights are used to discard pairs of symbols that are likely to arise from common source artifacts (such as blink artifacts or volume conduction). We set the values of *k* to 3 and *τ* to 16 ms, following procedures from previous single-word processing experiments (Hesse et al., 2019; García et al., 2020). This value of *τ* is sensitized to frequencies lower than 10 Hz, a range associated with lexico-semantic processes in bilinguals (Pérez et al., 2015; Vilas et al., 2019), including word translation (Grabner et al., 2007; Dottori et al., 2020). As for power analyses, we considered the wSMI patterns elicited in an early (0-300 ms) and a later (300-600 ms) time window. A detailed description of wSMI can be found in previous work (King et al., 2013).

As for frequency analyses, cluster-based statistics were performed on the wSMI topographical results. We ran a permutation test on the wSMI coefficient matrices obtained for each subject in each condition, to obtain clusters of connections based on neighboring criteria (measured through Euclidean distance). Two connections were considered neighbors if both of the electrodes in one connection were neighbors of the electrodes in the other connection. As cluster-level statistics, we used the largest sum of the *t*-values of all the connections forming the clusters observed in each permutation. To define connections that are candidates to form

significant clusters we performed two-tailed t-tests at $p_{con} < .05$. Significant clusters were then established at $p_{clus} < .05$. This cluster-based approach to whole-scalp connectivity solves the multiple comparisons problem and captures robust effects during language processing across varied populations, including bilinguals (Birba et al., 2020; Birba, 2021), circumventing the potential biases of estimating average connectivity between pre-defined ROIs (Dottori et al., 2017). As in previous research (Birba et al., 2020), the p-value of each cluster was estimated as the proportion of 2,000 random permutation of the wSMI matrices that yielded a cluster-level statistic greater than the sum of t-values of the corresponding cluster in the observed data. As for spectral power analyses, statistical comparisons were performed between FT and BT (to capture power signatures of the translation asymmetry effect) and between L2R and L1R (to rule out potential differences in relevant single-language processes).

Correlation between behavioral and electrophysiological measures

In line with reported approaches of the same tasks (Dottori et al., 2020), we performed correlations between mean RT and electrophysiological signatures (spectral power and connectivity) of each comparison yielding significant differences between conditions. Given that data were normally distributed (see below), we used Pearson's coefficient. As in previous reports of the same tasks (Dottori et al., 2020), correlations were considered significant if they yielded a $p < .05$ after an FDR correction for comparisons between frequency bands and time windows (Benjamini and Hochberg, 1995).

RESULTS

Behavioral results

Accuracy results (Table 1) revealed a main effect of task [$F_{1,26} = 135.92$; $p < .001$; $\eta_p^2 = 0.84$], with better performance on reading (98.85%) than translation (86.7%). The main effect of language was non-significant [$F_{1,26} = 1.09$; $p = .31$; $\eta_p^2 = 0.04$], and so was the interaction between task and language [$F_{1,26} = 0.02$; $p = .88$; $\eta_p^2 < .001$]. For details, see Supplementary material 2, Table S2.

RT results (Table 1, Figure 1B) showed a main effect of task [$F_{1,26,} = 96.30$; $p < .001$; $\eta_p^2 = 0.87$], with slower performance for translation (635.41) than reading (376.02 ms). The main effect of language was not significant [$F_{1,26} = 3.39$; $p = .08$; $\eta_p^2 = 0.12$]. However, a significant interaction emerged between task and language [$F_{1,26} = 28.09$; $p < .001$; $\eta_p^2 = 0.62$]. Post-hoc contrasts, corrected for multiple comparisons via Tukey's HSD test, showed that FT was significantly slower than any other task, BT was slower than any reading task, and L2R

9

was slower than L1R (all *p*-values < .05). For details, see Supplementary material 2, tables S3 and S4.

Accuracy and RT results were replicated upon running linear mixed-effects models, considering task and language as fixed effects, and subject as random effect. For details, see Supplementary material 3 (tables S5, S6, and S7).

Table 1. Descriptive statistics for accuracy and RT results.

Task	Mean accuracy (*SD*)	Mean RT (*SD*)
FT	0.87 (0.06)	651.63 (148.47)
BT	0.87 (0.06)	619.19 (160.3)
L1R	0.99 (0.01)	409.02 (94.10)
L2R	0.99 (0.02)	343.03 (58.74)

Data presented as mean (*SD*). FT: forward translation (from L1 to L2), BT: backward translation (from L2 to L1), L2R: L2 (non-native language) reading, L1R: L1 (native language) reading. RT: response time.

Time-frequency results

While both translation tasks showed desynchronization relative to baseline activity, time-frequency analysis showed a significant ($p_{elec} < .05$; $p_{clus} < .05$) asymmetry effect (Figure 1C). Relative to BT, FT elicited higher power in the early time window (0-300 ms) in the theta band (4-7 Hz) over a centro-frontal cluster ($p_{clus} = .01$) and lower power in the later time window (300-600 ms) in the beta frequency range, over a left centro-posterior cluster in the lower-beta range (14-20 Hz) ($p_{clus} = .02$) and a centro-frontal cluster in the upper-beta range (21-30 Hz) ($p_{clus} = .01$). No significant cluster was observed in the comparison between L1R and L2R for any combination of frequency band and time window.

Functional connectivity results

FT involved lower wSMI connectivity than BT in the early window (0-300 ms) over a distributed cluster spanning left frontal and right parietal electrodes ($p_{con} < .05$; $p_{clus} = .046$) – Figure 1D. No significant differences were observed between these tasks in the later window (300-600 ms). No significant differences were observed between L1R and L2R in either window.

Correlation results

To examine associations between behavioral and neurophysiological signatures of the translation asymmetry effect, we correlated the mean RT difference between FT and BT with the same subtraction over significant power and connectivity clusters. We used Pearson's correlations given that data was normally distributed (RT: Shapiro Wilk's test, $p = .25$; power: Shapiro Wilk's test, $p = .58$; functional connectivity: Shapiro Wilk's test, $p = .12$). For power analyses, a significant positive correlation was observed between RT and the theta-band cluster ($rho = 0.46$, $p = .03$), but not for the lower-beta ($rho = -0.36$; $p = .09$) or the upper-beta ($rho = 0.14$; $p = .49$) clusters. For functional connectivity analyses, the correlation between RT and wSMI over the significant fronto-parietal cluster was also non-significant ($rho = 0.24$; $p = .23$). See Figure 1E.

[INSERT FIGURE 1 HERE]

DISCUSSION

This study examined behavioral and neurophysiological signatures of translation asymmetry. Relative to BT, FT was characterized by slower RTs, differential (de)synchronization patterns in the theta and beta bands, and lower functional connectivity in frequencies below 10 Hz. Moreover, the RT difference between translation directions correlated with its spectral power signatures in the theta band. These findings afford new neural constraints for mainstream models of bilingual memory.

RT results showed that FT was slower than every other condition, crucially including BT. This replicates previous reports targeting the same language pair (Sáchez-Casas et al., 1992; Francis et al., 2014; García, 2019) as well as others, such as English-German (Kroll et al., 1994), Russian-English (Ibrahim et al., 2017), Dutch-English (Degroot et al., 1994; De Groot and Poot, 1997; Poarch et al., 2015), and Chinese-English (Cheung et al., 1998). Interestingly, our results came from a highly proficient bilingual group, which might seem surprising considering claims that translation asymmetry attenuates as L2 competence increases (McElree et al., 2000; Christoffels et al., 2006; García et al., 2014). However, slower performance in FT than BT has been reported in high-proficiency bilinguals and even in professional simultaneous interpreters (Santilli et al., 2019). Together with these findings, our

11

study suggests that, at least under certain circumstances, the translation asymmetry effect is pervasive enough to emerge even in persons with elevated L2 skills.

This was corroborated by oscillatory measures. Compared with BT, FT yielded greater fronto-central theta power in an early (0-300 ms) window. Power increases in the theta band have been shown to index greater lexico-semantic demands, including word retrieval and monitoring functions (Bastiaansen et al., 2005; Hald et al., 2006; Davidson and Indefrey, 2007; Kielar et al., 2014). More particularly, such modulations have also been linked to increased difficulty during translation, as theta power proved greater for low- vs. high-frequency source words, suggesting more elaborate lexical search operations (Grabner et al., 2007). Our results indicate that theta-band synchronization may also constitute a signature of the translation asymmetry effect, arguably reflecting higher lexical access and search demands for FT than BT. Of note, a previous study (Dottori et al., 2020) found that professional interpreters exhibited *lower* theta synchrony than non-interpreters during translation tasks. This further highlights the crucial role of theta modulations during word translation, suggesting that their synchronization and desynchronization might differentially index task- and expertise-related effects, respectively.

FT also elicited less spectral power than BT over the low- and high-beta bands in a late window (300-600 ms), across centro-posterior and centro-frontal electrodes, respectively. Beta desynchronization has been associated with the processing of semantically complex (open-class) vis-à-vis semantically simpler (closed-class) words (Bastiaansen et al., 2005), and with the processing of semantic violations –an affect attributed to heightened attentional demands during linguistic processing (Kielar et al., 2014). The beta-band signatures observed here might also reflect attentional demands for FT. Indeed, previous research has revealed greater P200 modulations (a core signature of attentional allocation) for FT than BT (Christoffels et al., 2013). Moreover, beta desynchronization has been linked to word translation proper, specifically in a late (> 400-ms) window (Grabner et al., 2007). In line with this antecedent, our combined theta- and beta-band results suggest that translation asymmetry is indexed by distinct coupling and decoupling dynamics over specific frequency bands as underlying processes unfold in time.

In the same vein, functional connectivity results revealed lower wSMI values for FT than BT. This pattern was observed in an early window (0-300 ms), over left frontal and right parietal electrodes, for frequencies below 10 Hz. Previously reported wSMI signatures of FT relative to BT (including reduced intracranial connectivity across temporal, frontal, and prefrontal areas) have been linked to greater attentional demands for the former direction

12

(García et al., 2016). Similar conclusions stem from fMRI connectivity results (lower temporo-thalamic and higher fronto-temporo-parietal connectivity for FT compared with BT), indicating reduced reliance on automatic relay mechanisms and a greater interplay of attentional and lexico-semantic processes (Zheng et al., 2020). Compatibly, our results reinforce the view that translation asymmetry may be characterized by distinct *integration* efforts across linguistic and executive systems. Moreover, they indicate that the neural basis of this effect involves not only transient (oscillatory) (de)coupling dynamics, but also coordinated activity patterns across segregated neural locations.

At least some of these electrophysiological effects seem directly related to the outward manifestation of translation asymmetry. The mean RT difference between translation tasks positively correlated with the corresponding early theta power difference. To our knowledge, this is the first demonstration of an association between behavioral and neural signatures of the effect. Notably, this correlation was selective for theta band modulations, which have already been associated with RT measures during BT and FT tasks separately (Dottori et al., 2020). These results further highlight the crucial role of theta oscillations as a critical marker of word translation efficiency, in general, and its sensitivity to directionality, in particular.

Our findings carry theoretical implications. Previous neuroscientific evidence on translation asymmetry was restricted to hemodynamic patterns (Klein et al., 1995; Rinne et al., 2000; Tommola et al., 2000) and event-related potentials (Christoffels et al., 2013). The present study indicates that this effect is also related to (de)synchronization patterns that capture topographically sparse modulations. Moreover, all the observed EEG patterns emerged in the absence of differences between reading tasks. Thus, the reported neural signatures of translation asymmetry were likely not primarily driven by more basic processes implied by word translation (e.g., single-word reading and production), but rather by translation-specific dynamics –arguably, those involved in cross-linguistic operations proper. In this sense, they may reflect the stage mediating source-item input and target-item output –a critical phase captured in diverse models of interlingual reformulation (Seleskovitch, 1968; Seleskovitch and Lederer, 1984; Bell and Candlin, 1991; Marianne, 1994; García, 2019).

More generally, our study bridges the gap between psycholinguistic and neuroscientific conceptualizations of translation asymmetry. Pioneering and recent accounts, such as the Revised Hierarchical Model (Kroll et al., 1994) and Multilink (Dijkstra et al., 2019), were forged exclusively on behavioral data, leading to explanations based on differential "connection strengths" between lexical and conceptual systems, without any biological grounding. On the other hand, classical studies (Rinne et al., 2000; Tommola et al., 2000;

13

Quaresima et al., 2002) and different models (Fabbro, 1999; García, 2019) of neural dissociations between FT and BT failed to consider RTs or temporally precise brain metrics, limiting the interpretability of results. Our combined findings capture the multidimensional nature of the effect and the interplay among its signatures, inviting integrative reconceptualizations of a distinguishing trait of bilingualism.

Our study is not without limitations. First, although our sample was large enough to reach adequate effect sizes and surpassed those of other studies (Grabner et al., 2007; Kielar et al., 2014; Elmer and Kühnis, 2016), more robust results could be obtained with more participants. Second, participants were predominantly female. Although diverse (neuro)linguistic effects prove similar between with (Ibrahim et al., 2017) and without such an imbalance (Aravena et al., 2010; Poarch et al., 2015), and even though systematic reviews (Wallentin, 2009) and meta-analyses (Sato, 2020) reveal little to no sex-related effects during language processing, future studies should replicate our work with better balanced samples. Third, all frequency and functional connectivity analyses were performed over previously defined time windows, following previous neurolinguistic studies (Hald et al., 2006; Grabner et al., 2007; Willems et al., 2008; Vilas et al., 2019). Moreover, the frequency ranges for which these analyses were sensitized were also determined *a priori*, based on previously reported procedures (Hesse et al., 2019; Dottori et al., 2020; García et al., 2020). While this approach favors comparability with relevant works, it would be interesting to test whether similar results emerge from data-driven approaches. Fourth, other functional connectivity measures could be used to test the specificity of our wSMI results. Finally, replications would also be desirable across different language pairs and even considering subgroups with different levels of L2 competence, ages of L2 appropriation, or formal translation training.

In conclusion, we showed that translation asymmetry is underpinned by fast-changing oscillatory and functional connectivity signatures, with theta decoupling emerging as a direct correlate of RT outcomes. The differential routes posited for FT and BT in mainstream bilingual memory models thus seem critically subserved by highly specific neural (de)synchronization patterns. By integrating multidimensional measures of the effect and capturing direct associations between them, our study paves the way for more comprehensive accounts of a key feature of bilingual lexical processing.

GLOSSARY

- BT: backward translation (from L2 to L1).
- EEG: electroencephalography (a research technique capturing neurophysiological changes associated with specific cognitive processes).
- fMRI: functional magnetic resonance imaging (a neuroimaging technique revealing hemodynamic changes associated with particular cognitive processes).
- FT: forward translation (from L1 to L2).
- Functional connectivity: linear or nonlinear covariations between brain activity fluctuations in different recording sites.
- L1: native language.
- L1R: reading in L1.
- L2: non-native language.
- L2R: reading in L2.
- RT: response time.
- Translation asymmetry effect: the detection of different cognitive demands for FT and BT, typically manifested as lower accuracy and/or longer RT for the former task.
- wSMI: weighted Symbolic Mutual Information (a non-linear functional connectivity metric).

15

does not represent the official views of the National Institutes of Health, Alzheimer's Association, Rainwater Charitable Foundation, or Global Brain Health Institute.

DECLARATION OF INTEREST

None to report.

REFERENCES

Abdi H, Williams L (2010) Tukey's honestly significant difference (HSD) test. In: Encyclopedia of research design, vol. 3 (Salkind NJ, ed), pp 1-5. Thousand Oaks, California: SAGE Publications.

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

Aravena P, Hurtado E, Riveros R, Cardona JF, Manes F, Ibáñez A (2010) Applauding with closed hands: neural signature of action-sentence compatibility effects. PLoS ONE 5:e11751-e11751.

Bastiaansen M, Magyari L, Hagoort P (2010) Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. J Cogn Neurosci 22(7):1333-1347.

Bastiaansen MCM, Linden Mvd, Keurs Mt, Dijkstra T, Hagoort P (2005) Theta responses are involved in lexical—semantic retrieval during language processing. J Cogn Neurosci 17(3):530-541.

Bell RT (1991) Translation and translating: theory and practice (Candlin CN, ed). London: Longman.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57(1):289-300.

Bialystok E, Craik FI, Luk G (2012) Bilingualism: consequences for mind and brain. Trends Cogn Sci 16(4):240-250.

Birba A, Beltrán D, Martorell Caro M, Trevisan P, Kogan B, Sedeño L, Ibáñez A, García AM (2020) Motor-system dynamics during naturalistic reading of action narratives in first and second language. Neuroimage 216:116820.

Birba A, Fittipaldi S, Cediel Escobar J, Gonzalez Campo C, Legaz A, Galiani A, Díaz Rivera M, Martorell Caro M, et al. (accepted) Multidimensional neurocognitive markers of ecological discourse typify diverse neurodegenerative diseases. Cereb Cortex.

Buzsaki G (2006) Rhythms of the brain. New York: Oxford University Press.

Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. Science 304(5679):1926-1929. https://doi: 10.1126/science.1099745.

Cheung H, Chen H-C (1998) Lexical and conceptual processing in Chinese-English bilinguals: further evidence for asymmetry. Mem Cognit 26(5):1002-1013.

Christoffels IK, De Groot AM, Kroll JF (2006) Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. J Mem Lang 54(3):324-345.

Christoffels IK, Ganushchak L, Koester D (2013) Language conflict in translation: an ERP study of translation production. J Cogn Psychol 25(5):646-664.

Davidson DJ, Indefrey P (2007) An inverse relation between event-related and time–frequency violation responses in sentence processing. Brain Res J 1158:81-92.

Davies M (2002) Corpus del español: 100 million words, 1200s-1900s.

Davies M (2010) The corpus of contemporary American English as the first reliable monitor corpus of English. Lit Linguistics Comput 25(4):447-464.

De Groot AM, Poot R (1997) Word translation at three levels of proficiency in a second language: the ubiquitous involvement of conceptual memory. Lang Learn 47(2):215-264.

Degroot AM, Dannenburg L, Vanhell JG (1994) Forward and backward word translation by bilinguals. J Mem Lang 33(5):600-629.

Dijkstra A, Van Heuven WJ (2002) The architecture of the bilingual word recognition system: from identification to decision. Biling (Camb Engl) 5(3):175-197.

Dijkstra T, Van Heuven WJ (1998) The BIA model and bilingual word recognition. In: Localist connectionist approaches to human cognition, (Grainger J, Jacobs AM, eds), pp 189-225. Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.

Dijkstra T, Wahl A, Buytenhuijs F, Van Halem N, Al-Jibouri Z, De Korte M, Rekké S (2019) Multilink: a computational model for bilingual word recognition and word translation. Biling (Camb Engl) 22(4):657-679.

Dottori M, Hesse E, Santilli M, Vilas MG, Martorell Caro M, Fraiman D, Sedeño L, Ibáñez A, et al. (2020) Task-specific signatures in the expert brain: differential correlates of translation and reading in professional interpreters. Neuroimage 209:116519.

Dottori M, Sedeño L, Martorell Caro M, Alifano F, Hesse E, Mikulan E, García AM, Ruiz-Tagle A, et al. (2017) Towards affordable biomarkers of frontotemporal dementia: a classification study via network's information sharing. Sci Rep 7(1):3822.

Elmer S, Kühnis J (2016) Functional connectivity in the left dorsal stream facilitates simultaneous language translation: an EEG study. Front Hum Neurosci 10:60.

Fabbro F (1999) The neurolinguistics of bilingualism: an introduction. Hove: Psychology Press.

Fan X, Wu Y, Cai L, Ma J, Pan N, Xu X, Sun T, Jing J, et al. (2021) The differences in the whole-brain functional network between Cantonese-Mandarin bilinguals and Mandarin monolinguals. Brain Sci 11(3):310

Francis WS, Tokowicz N, Kroll JF (2014) The consequences of language proficiency and difficulty of lexical access for translation performance and priming. Mem Cognit 42(1):27-40.

French RM, Jacquet M (2004) Understanding bilingual memory: models and data. Trends Cogn Sci 8(2):87-93.

Friston KJ (2011) Functional and effective connectivity: a review. Brain Connect 1(1):13-36.

19

García AM (2015) Translating with an injured brain: neurolinguistic aspects of translation as revealed by bilinguals with cerebral lesions. Meta 60(1):112-134.

García AM, Mikulan E, Ibáñez A (2016) A neuroscientific toolkit for translation studies. In: Reembedding Translation Process Research, vol. 128 (Muñoz Martín R, ed), pp 21-46. Amsterdam/Philadelphia: John Benjamins.

García AM (2019) The neurocognition of translation and interpreting, (Valderón RA, ed). Amsterdam/Philadelphia: John Benjamins.

García AM, Hesse E, Birba A, Adolfi F, Mikulan E, Martorell Caro M, Petroni A, Bekinschtein TA, et al. (2020) Time to face language: embodied mechanisms underpin the inception of face-related meanings in the human brain. Cereb Cortex 30(11):6051-6068. doi: https://doi.org/10.1093/cercor/bhaa178.

García AM, Ibáñez A, Huepe D, Houck AL, Michon M, Lezama CG, Chadha S, Rivera-Rei A (2014) Word reading and translation in bilinguals: the impact of formal and informal translation expertise. Front Psychol 5:1302-1302. doi: https://doi.org/10.3389/fpsyg.2014.01302.

García AM (2013) Brain activity during translation: a review of the neuroimaging evidence as a testing ground for clinically-based hypotheses. J Neurolinguistics 26(3):370-383.

Grabner RH, 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.

Grady CL, Luk G, Craik FI, Bialystok E (2015) Brain network activity in monolingual and bilingual older adults. Neuropsychologia 66:170-181.

Grundy JG, Anderson JAE, Bialystok E (2017a) Bilinguals have more complex EEG brain signals in occipital regions than monolinguals. Neuroimage 159:280-288.

Grundy JG, Anderson JAE, Bialystok E (2017b) Neural correlates of cognitive processing in monolinguals and bilinguals. Ann N Y Acad Sci 1396(1):183-201.

Hald LA, Bastiaansen MC, Hagoort P (2006) EEG theta and gamma responses to semantic violations in online sentence processing. Brain Lang 96(1):90-105.

Harris B, Sherwood B (1978) Translating as an innate skill. In: Language interpretation and communication, vol. 6 (Gerver D, Sinaiko HW, eds), pp. 155-170. Boston, Massachusetts: Springer.

Hesse E, Mikulan E, Sitt JD, del Carmen García M, Silva W, Ciraolo C, Vaucheret E, Raimondo F, et al. (2019) Consistent gradient of performance and decoding of stimulus type and valence from local and network activity. IEEE Trans Neural Syst Rehabil Eng 27(4):619-629.

Ibrahim A, Cowell PE, Varley RA (2017) Word frequency predicts translation asymmetry. J Mem Lang 95:49-67.

Kielar A, Meltzer JA, Moreno S, Alain C, Bialystok E (2014) Oscillatory responses to semantic and syntactic violations. J Cogn Neurosci 26(12):2840-2862.

King JR, Sitt JD, Faugeras F, Rohaut B, El Karoui I, Cohen L, Naccache L, Dehaene S (2013) Information sharing in the brain indexes consciousness in noncommunicative patients. Curr Biol 23(19):1914-1919.

Klein D, Milner B, Zatorre RJ, Meyer E, Evans AC (1995) The neural substrates underlying word generation: a bilingual functional-imaging study. Proc Natl Acad Sci U S A 92(7):2899.

Kroll JF, Gullifer JW, Rossi E (2013) The multilingual lexicon: the cognitive and neural basis of lexical comprehension and production in two or more languages. Annu Rev Appl Linguist 33:102-127.

21

Kroll JF, Stewart E (1994) Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. J Mem Lang 33(2):149.

Kroll JF, Van Hell JG, Tokowicz N, Green DW (2010), The revised hierarchical model: a critical review and assessment. Biling (Camb Engl) 13(3):373.

Litcofsky KA, Van Hell JG (2017), Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. Neuropsychologia 97:112-139.

Love J, Selker R, Marsman M, Jamil T, Dropmann D, Verhagen J, Ly A, Gronau QF, et al. (2019) JASP: graphical statistical software for common statistical designs. J Stat Softw 88(1):1-17.

Malakoff ME (1992) Translation ability: a natural bilingual and metalinguistic skill. In: Advances in psychology, vol. 83 (Burns B, ed), pp 515-529. North-Holland: Elsevier.

Marianne L, (1994) La traduction aujourd'hui. Le modèle interprétatif. Paris: Hachette.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164(1):177-190.

McElree B, Jia G, Litvak A (2000) The time course of conceptual processing in three bilingual populations. J Mem Lang 42(2):229-254.

Mišić B, Sporns O (2016) From regions to connections and networks: new bridges between brain and behavior. Curr Opin Neurobiol 40:1-7.

Moguilner S, Birba A, Fino D, Isoardi R, Huetagoyena C, Otoya R, Tirapu V, Cremaschi F, et al. (2021) Multimodal neurocognitive markers of frontal lobe epilepsy: insights from ecological text processing. Neuroimage 235:117998.

Pérez A, Dowens MG, Molinaro N, Iturria-Medina Y, Barraza P, García-Pentón L, Carreiras M (2015) Complex brain network properties in late L2 learners and native speakers. Neuropsychologia 68:209-217.

Poarch GJ, Van Hell JG, Kroll JF (2015) Accessing word meaning in beginning second language learners: lexical or conceptual mediation? Biling (Camb Engl) 18(3):357-371.

Quaresima V, Ferrari M, van der Sluijs MC, Menssen J, Colier WN (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.

Rinne JO, Tommola J, Laine M, Krause BJ, Schmidt D, Kaasinen V, Teräs M, Sipilä H, et al. (2000) The translating brain: cerebral activation patterns during simultaneous interpreting. Neurosci Lett 294(2):85-88.

Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52(3):1059-1069.

Rubinov M, Sporns O (2011) Weight-conserving characterization of complex functional brain networks. Neuroimage 56(4):2068-2079.

Sáchez-Casas RM, García-Albea JE, Davis CW (1992) Bilingual lexical processing: exploring the cognate/non-cognate distinction. Eur J Cogn Psychol 4(4):293-310.

Santilli M, Vilas MG, Mikulan E, Martorell Caro M, Muñoz E, Sedeño L, Ibáñez A, García AM (2019) Bilingual memory, to the extreme: lexical processing in simultaneous interpreters. Biling (Camb Engl) 22(2):331-348.

Sato M (2020) The neurobiology of sex differences during language processing in healthy adults: a systematic review and a meta-analysis. Neuropsychologia 140:107404.

Seleskovitch D (1968) L'interprète dans les conférences internationales. Problèmes de langage et de communication. Paris: Minard.

Seleskovitch D, Lederer M (1984) Interpréter pour traduire. In: L'Information Grammaticale,

    vol. 25, pp 44-47. Paris: Didier érudition.

Sulpizio S, Del Maschio N, Del Mauro G, Fedeli D, Abutalebi J (2020) Bilingualism as a

    gradient measure modulates functional connectivity of language and control networks.

    Neuroimage 205:116306.

Tommola J, Laine M, Sunnari M, Rinne JO (2000) Images of shadowing and interpreting.

Interpreting (Amst) 5(2):147-167.

Vilas MG, Santilli M, Mikulan E, Adolfi F, Martorell Caro M, Manes F, Herrera E, Sedeño L,

    et al. (2019) Reading Shakespearean tropes in a foreign tongue: age of L2 acquisition

    modulates neural responses to functional shifts. Neuropsychologia 124:79-86.

Wallentin M (2009) Putative sex differences in verbal abilities and language cortex: a critical

    review. Brain Lang 108(3):175-183.

Weiss S, Mueller HM, Schack B, King JW, Kutas M, Rappelsberger P (2005) Increased

    neuronal communication accompanying sentence comprehension. Int J Psychophysiol

    57(2):129-141.

Willems RM, 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.

Zheng B, Báez S, Su L, Xiang X, Weis S, Ibáñez A, García AM (2020) Semantic and

    attentional networks in bilingual processing: fMRI connectivity signatures of

    translation directionality. Brain Cogn 143:105584.

Figure 1
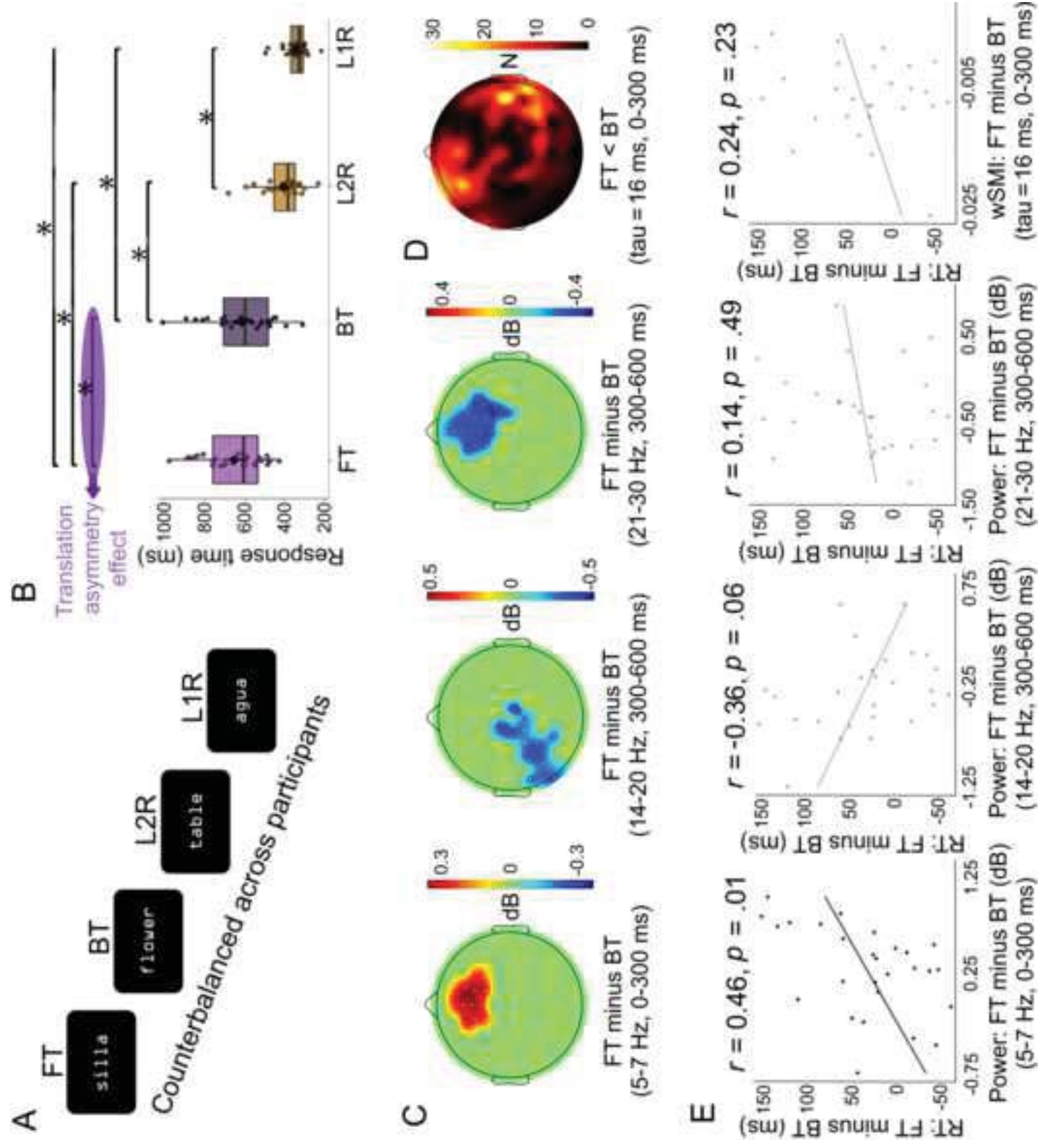
Figure 1 caption

Figure caption

Figure 1. Experimental conditions and results. **A.** The experiment comprised four counterbalanced blocks of 64 stimuli, two in L1 (for the FT and L1R tasks) and two in L2 (for the BT and the L2R tasks). **B.** Response times for each experimental condition. FT was slower than every other condition, crucially including BT. The asterisk (*) indicates significant differences. **C.** Clusters yielding significantly different power between conditions. Compared to BT, FT elicited greater theta power in an early window (0-300 ms) and reduced lower-beta and upper-beta power in a late time window (300-600 ms). Each plot shows the mean difference in power (baseline-normalized and converted to dB) between the corresponding experimental conditions. **D.** Cluster yielding significantly different connectivity between conditions. FT yielded lower wSMI values than BT in an early window, in frequencies below 10 Hz. The colorbar indicates the number of connections between (a) electrodes in the cluster and (b) every other electrode in the scalp. **E.** Scatterplots of the associations between the mean power difference (first three insets) and mean wSMI difference (fourth inset) of significant clusters and the corresponding mean response time difference between FT and BT. A significant correlation was found only for the early theta cluster. Every other association was non-significant. BT: backward translation; FT: forward translation; L1R: native-language reading; L2R: foreign-language reading. RT: response time; dB: decibels.

Supplementary material

1. Power estimation

Using G*Power 3.1 (Faul, 2007), we estimated power for a repeated measures ANOVA, with $p = .05$, $\eta_p^2 = 0.13$ –based on previous results from the same task (Dottori et al., 2020)–, and power $= 0.95$. Results indicated that a total sample size of 16 was enough to reach the estimated effects. Our actual sample size ($n = 27$) reaches a power of .99.

2. Supplementary ANOVA results

Table S1. Comparison of rejected trials per condition for response time analyses.

Effect	F	p-value*	η_p^2
Task	43.08	< .001	0.62
Language	1.82	.19	0.07
Task*Language	0.73	.40	0.03

*p-values calculated with a 2x2 repeated measures ANOVA.

Table S2. Accuracy results based on a repeated measures ANOVA.

Effect	F	p-value*	η_p^2
Task	135.92	< .001	0.84
Language	1.09	.31	0.04
Task*Language	0.02	.88	< 0.001

*p-values calculated with a 2x2 repeated measures ANOVA.

Table S3. Response time results based on a repeated measures ANOVA.

Effect	F	p-value*	η_p^2
Task	165.10	< .001	0.86
Language	3.39	.08	0.12
Task*Language	42.79	< .001	0.62

*p-values calculated with a 2x2 repeated measures ANOVA.

Table S4. Post-hoc analyses of response time results.

Pairwise contrast		t-statistic	p_{Tukey}	Cohen's d
FT	L1R	14.33	< .001	2.49
	L2R	10.95	< .001	2.76
	BT	2.75	.04	0.52
BT	L1R	12.47	< .001	2.02
	L2R	9.76	< .001	2.13
L2R	L1R	5.59	< .001	1.08

3. Supplementary linear mixed effect model results

Table S5. Accuracy results in a linear mixed-effects model.

Effect	F	p-value*
Task	131.35	< .001
Language	1.08	.30
Task*Language	0.03	.86

*p-values calculated with a linear mixed-effects model.

Table S6. Response time results in a linear mixed-effects model.

Effect	F	p-value*
Task	165.09	< .001
Language	3.29	.08
Task*Language	44.85	< .001

*p-values calculated with a linear mixed-effects model.

Table S7. Pairwise comparisons of RT data in a linear mixed-effects model.

Pairwise contrast		t-statistic	p_{Tukey}
FT	L1R	14.36	< .001
	L2R	14.24	< .001
	BT	2.75	.049
BT	L1R	10.47	< .001
	L2R	9.78	< .001
L2R	L1R	5.59	< .001

Supplementary references

Dottori M, Hesse E, Santilli M, Vilas M, Martorell Caro M, Fraiman D, Sedeño L, Ibáñez A & García AM (2020). Task-specific signatures in the expert brain: Differential correlates of translation and reading in professional interpreters. *NeuroImage* 209: 116519.

Faul FE, Lang E, Buchner A (2007), G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39: 175-191.

Ethical statement

All participants signed an informed consent, and all experimental protocols were performed in accordance with the Declaration of Helsinki. This study was approved by the institutional ethics committee.