Brain and Cognition 134 (2019) 122–134

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Neuroplasticity, bilingualism, and mental mathematics: A behavior-MEG study

Jo-Fu Lotus Lin[a,b,c], Toshiaki Imada[a,b], Patricia K. Kuhl[a,b,*]

[a] Institute for Learning & Brain Sciences, University of Washington, Seattle, WA, USA
[b] Department of Speech & Hearing Sciences, University of Washington, Seattle, WA, USA
[c] Institute of Linguistics, National Tsing Hua University, Taiwan

ARTICLE INFO

Keywords:
Bilinguals
Mental calculation
Magnetoencephalography (MEG)

ABSTRACT

Bilingual experience alters brain structure and enhances certain cognitive functions. Bilingualism can also affect mathematical processing. Reduced accuracy is commonly reported when arithmetic problems are presented in bilinguals' second (L2) vs. first (L1) language. We used MEG brain imaging during mental addition to characterize spatiotemporal dynamics during mental addition in bilingual adults. Numbers were presented auditorally and sequentially in bilinguals' L1 and L2, and brain and behavioral data were collected simultaneously. Behaviorally, bilinguals showed lower accuracy for two-digit addition in L2 compared to L1. Brain data showed stronger response magnitude in L2 versus L1 prior to calculation, especially when two-digit numbers were involved. Brain and behavioral data were significantly correlated. Taken together, our results suggest that differences between languages emerge prior to mathematical calculation, with implications for the role of language in mathematics.

1. Introduction

1.1. Bilingual experience, connectivity, and cognition

Numerous studies now show that bilingualism alters both brain and behavior. Significant connectivity differences measured with diffusion tensor imaging (DTI), a neuroimaging tool for measuring characteristics of white matter, have been reported between monolingual and bilingual brains (for a review, see García-Pentón, Garcia, Costello, Duñabeitia, & Carrieras, 2016). DTI measures are sensitive to white matter structural features and a feature of DTI, fractional anisotropy (FA), which measures the directional asymmetry of the diffusion of water molecules, differs between monolingual and bilingual children (e.g., Mohades et al., 2015), adults (e.g., Pliatsikas, Moschopoulou, & Saddy, 2015; Kuhl et al., 2016), and aging populations (Gold, Johnson, & Powell, 2013; Luk, Bialystok, Craik, & Grady, 2011). Differences between the two populations extend to cognition.

Cognitive enhancements, such as those shown in executive function skills, have been consistently reported for bilinguals throughout the lifespan. Infants from bilingual families learn speech structures more easily (Kovacs & Mehler, 2009a, 2009b), and bilingual toddlers show more flexibility when interpreting word forms (Estes & Hay, 2015). In bilingual children, advantages have been shown on metalinguistic tasks (Cromdal, 1999), switching tasks (Bialystok & Martin, 2004), Stroop tasks (Martin-Rhee & Bialystok, 2008), and theory of mind tasks (Goetz, 2003). Bilingual adults consistently outperform monolinguals on conflict resolution and Stroop tasks (Bialystok, Craik, & Luk, 2008, but also see Paap & Greenberg, 2013). These differences favoring bilinguals are thought to arise as a result of practice resolving conflict at the linguistic level between the two languages as bilinguals select or switch between languages and more generally manage attention between the two languages (Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001; Stocco & Prat, 2014).

Bilingual experience also impacts other aspects of cognition, such as arithmetic processing. When the language in which a bilingual's initial math facts were learned differs from the language used to test math facts in that person, lower performance accuracy and slower processing speed has been observed; this is the case both for early and late bilinguals (Grabner, Saalbach, & Eckstein, 2012; Saalbach, Eckstein, Andri, Hobi, & Grabner, 2013; Spelke & Tsivkin, 2001; Venkatraman, Siong, Chee, & Ansari, 2006). Even in early and balanced bilinguals, arithmetic facts are primarily learned in one language and that language optimizes mental calculation (Vaid & Menon, 2000). Cognitive costs related to switching between the language of encoding and the language of retrieval, have also been identified in domains other than arithmetic (Marian & Fausey, 2006; Marian & Neisser, 2000).

* Corresponding author at: Institute for Learning & Brain Sciences, University of Washington, Box 357988, Seattle, WA 98195-7988, USA.
E-mail address: pkkuhl@u.washington.edu (P.K. Kuhl).

https://doi.org/10.1016/j.bandc.2019.03.006
Received 28 February 2018; Received in revised form 22 July 2018; Accepted 28 March 2019
Available online 08 April 2019

1.2. Bilingual mental calculation

Behavioral studies show reduced accuracy when adult and child bilinguals are tested with arithmetic problems presented in their second or non-preferred language (L2) compared to their first or preferred language (L1) (Marsh & Maki, 1976; McClain & Huang, 1982; Murphey, 2014). To perform mental addition tasks, a sequence of cognitive processes are involved, which include encoding the numbers, lexical access, retrieval of arithmetic facts, maintaining results in memory, and responding with the arithmetic answer (Caramazza & McCloskey, 1987). Several models and theoretical frameworks have been proposed for understanding numerical and mathematical processing. Dehaene's Triple-Code Model (Dehaene & Cohen, 1995; Dehaene, 1992) assumes that numerical information is represented in three codes: the analog magnitude code, the visual Arabic code, and the verbal code. Each code serves difference functions. For example, the verbal codes are assumed to store arithmetic facts learned by rote memorization, such as addition and multiplication tables. Thus, when arithmetic problems are presented in Arabic digits, transcoding from Arabic digits to verbal codes must take place before retrieving arithmetic facts which are stored in verbal codes. However, the Triple-Code Model does not model verbal codes of more than one language. Thus, the model does not account for performance differences across languages.

Based on the Triple-Code Model, it is not clear whether arithmetic facts are stored only in the language they are taught. If arithmetic facts are stored only in the specific language in which arithmetic facts were initially learned, translating numbers into the specific language is required before retrieving arithmetic facts in the specific language. Thus, when the arithmetic problems are presented in the language different from the language in which arithmetic facts are stored, the chance of errors and interferences increases. This assumption is supported by behavioral studies showing lower performance levels when arithmetic problems were not presented in the language of learning arithmetic (Bernardo, 2001; Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Salillas & Wicha, 2012; Spelke & Tsivkin, 2001).

Campbell's Encoding Complex Model (Campbell & Clark, 1988; Campbell & Epp, 2004) describes verbal codes differently for L1 and L2. This model assumes that arithmetic facts are stored in each language separately. A variety of codes are included in the model: the magnitude code, the Arabic visual code, the L1 code, and the L2 code. These codes are connected in an associative network. The strength of each code and connection depends on previous experience with a particular task in a specific code. Thus, the Encoding Complex Model claims a strong link with languages used for learning arithmetic and retrieving arithmetic facts. This model also predicts that the strength of links can be increased with practices or familiarity. This prediction is supported by a recent ERP study (Martinez-Lincoln, Cortinas, & Wicha, 2015). Martinez-Lincoln et al. (2015) found that extensive experience using the other language, not the language of learning arithmetic, could mitigate the differences across languages.

Based on the Triple-Code Model, arithmetic facts are stored in and retrieved by verbal codes (Dehaene & Cohen, 1995; Dehaene, 1992). According to the Encoding Complex Model (Campbell & Clark, 1988; Campbell & Epp, 2004), the connection strength is stronger between L1 verbal codes and arithmetic facts. Therefore, to retrieve arithmetic codes stored in L1 verbal codes, a possible transcoding (into Arabic digits) and an additional translation (into L1 words) are required. For multi-digit addition, it is possible that bilinguals translate or transcode L1 and L2 numbers into other formats, such as Arabic digits. Based on Dehaene's Triple-Code model, multi-digit numbers are mediated via visual Arabic digit codes. However, according to the Triple-Code model, arithmetic facts are stored in the verbal code. It is only via the verbal code one can access to arithmetic facts. Thus, multi-digit addition would depend on the visual Arabic code as well as the verbal code.

Moreover, there is empirical evidence of translation. Bilinguals reported switching languages or translating numbers into their preferred language when performing arithmetic calculations (Moschkovich, 2007). Behavioral studies have attributed the reaction time differences between languages to the process of translating answers from the preferred to the non-preferred language (Marsh & Maki, 1976; McClain & Huang, 1982). Training studies on bilinguals have suggested bilinguals translated problems into the language of training or learning (Spelke & Tsivkin, 2001; Venkatraman et al., 2006).

1.3. Using brain measures to understand bilingual mental calculation

The underlying causes of reduced accuracy in L2 mathematical performance can be investigated using either behavioral or neuroimaging measures. Nonetheless, behavioral measures provide only a final outcome of a series of cognitive processes, and the temporal resolution provided by behavioral measures does not allow us to separate the translation and calculation process. However, neuroimaging techniques, such as magnetoencephalography (MEG), enable us to examine neural activity with millisecond precision. The fine temporal precision provided by MEG offers the opportunity to isolate cognitive processes that occur sequentially in time, separating the translation and calculation processes, and MEG's spatial resolution allows us to identify the sources of brain activities during mental addition.

Previous functional magnetic resonance imaging (fMRI) studies have reported activation from widespread cortical areas for all basic operations of mental arithmetic, including the inferior frontal, inferior parietal, and superior parietal areas during mental calculation (Arsalidou & Taylor, 2011; Dehaene, Molko, Cohen, & Wilson, 2004; Fehr, Code, & Herrmann, 2007; Grabner et al., 2007; Jost, Khader, Burke, Bien, & Rosler, 2009; Klein et al., 2016; Prado et al., 2011; Richard et al., 2000); however, fMRI has poor temporal resolution. Electrophysiological measures have excellent temporal resolution, but poor spatial resolution. Previous authors using event-related potential (ERP) components interpreted early components as reflecting attention processes, sensory and perceptual analysis of the stimuli, and identification of the numerals, whereas the later ERP components were assumed to reflect higher cognitive processes of arithmetic (Iguchi & Hashimoto, 2000; Jasinski & Coch, 2012; Jost, Khader, Burke, Bien, & Rosler, 2011; Nunez-Pena, Gracia-Bafalluy, & Tubau, 2011; Nunez-Pena, Honrubia-Serrano, & Escera, 2005; Prieto-Corona et al., 2010). Recent intracranial EEG measurements or electrocorticography (ECoG) also suggested later sustained activity in the parietal regions to be associated with arithmetic processing (Daitch et al., 2016; Dastjerdi, Ozker, Foster, Rangarajan, & Parvizi, 2013). However, cortical areas associated with calculation, such as the inferior frontal areas, are also found engaged during translation (Klein, Milner, Zatorre, Meyer, & Evans, 1995; Lehtonen et al., 2005; Rinne et al., 2000). Moreover, both calculation (Iguchi & Hashimoto, 2000; Nunez-Pena, Cortinas, & Escera, 2006) and translation (Phillips, Klein, Mercier, & de Boysson, 2006; Thierry & Wu, 2007) are associated with late event-related components. Thus, although event-related components provide fine temporal resolution, event-related components associated with calculation and translation can overlap if the study is not designed to separate them.

1.4. Current study

The goal of our study is to characterize the temporal dynamics during bilingual mental addition. To better illustrate differences between mental addition in L1 and L2, we constructed a schematic including hypothetical processes involved during two-digit plus one-digit addition in L2 (Fig. 1). This schematic also combines cognitive components with their associated cortical responses based on previous hemodynamic and electrophysiological studies.

We presented two-digit plus one-digit addition problems (*Complex Addition*, e.g., 62 + 9 = 71?) auditorally, resulting in presenting arithmetic numbers one by one sequentially, in participants' L1 vs. their

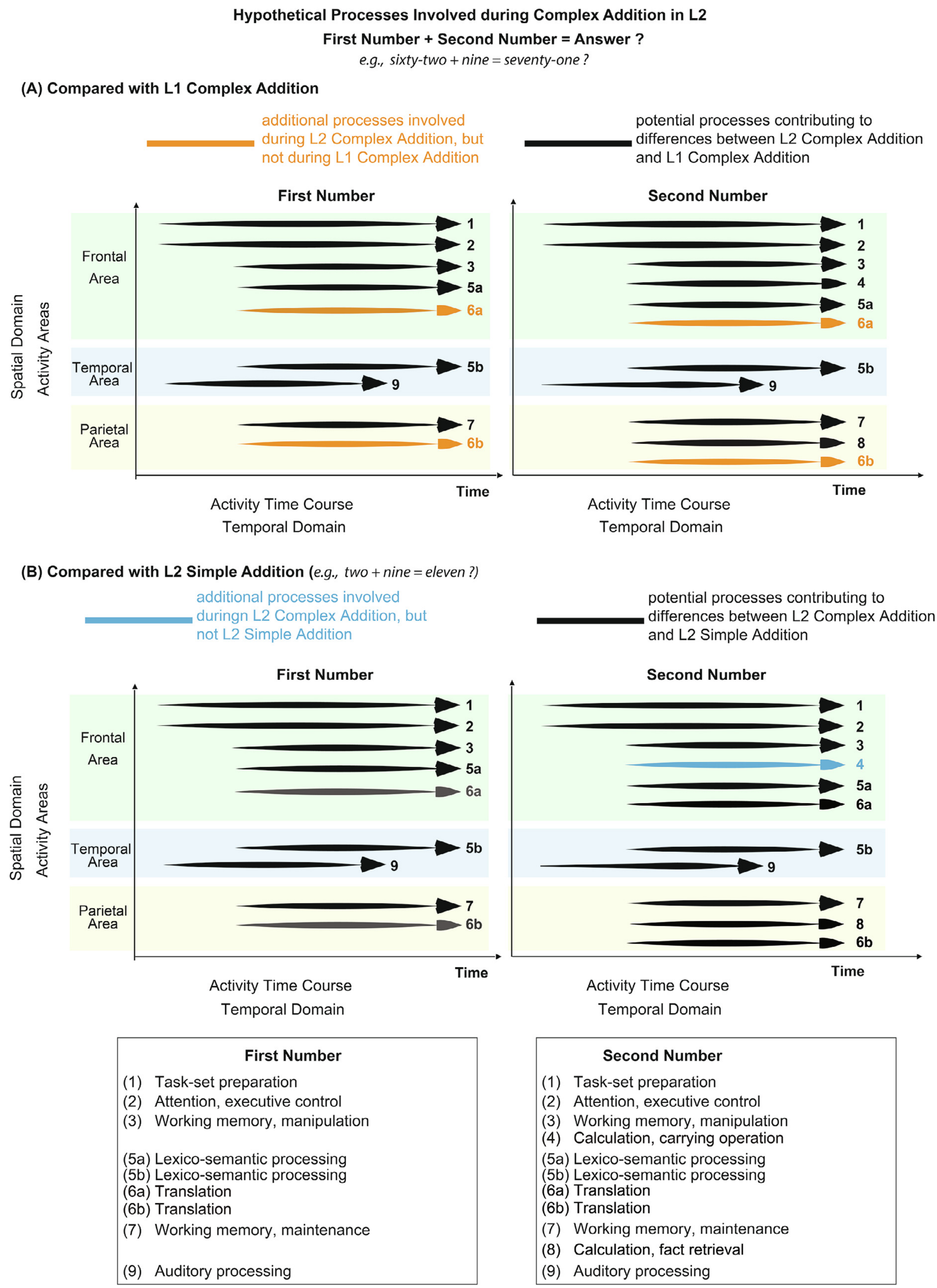

Fig. 1. A schematic incorporating cognitive processes and their associated spatiotemporal dynamics during mental addition. (a) A schematic comparing cognitive processes involved during L2 Complex Addition versus L1 Complex Addition. The additional cognitive processes involved in L2 Complex Addition but not in L1 Complex Addition are colored in orange. (b) A schematic comparing cognitive processes involved during L2 Complex Addition versus L2 Simple Addition. The additional cognitive processes involved in L2 Complex Addition but not in L2 Simple Addition are colored in aqua blue.

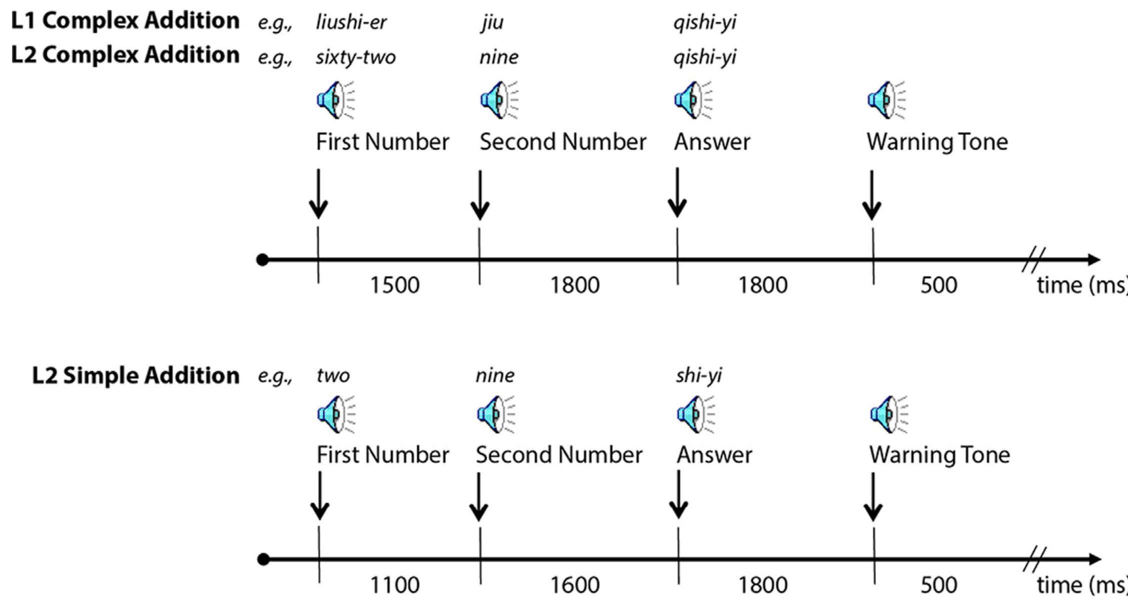

Fig. 2. Experimental Paradigm. Three tasks and example problems used in the present study. The three tasks include the L1 Complex Addition, L2 Complex Addition, and L2 Simple Addition. Participants were required to respond whether the provided answer was correct or incorrect before the Warning Tone. Timing of each interval (First Number, Second Number, and Answer) is shown. The arrows mark the stimulus onset of the auditorally presented numbers.

L2. Take 62 + 9 for example, we first presented '62' (First Number), then '9' (Second Number), and finally a correct or incorrect answer (Answer; '71' or '72') (Fig. 2). Participants were required to respond with a response key whether the Answer agreed with the sum of the First and Second Numbers. Addition problems were presented in participants' L1 or L2 in two separate sessions. The answers were always provided in the participants' L1.

Using MEG measures and a sequential presentation of numbers allow us to compare brain activity separately during the First Number and Second Number intervals of the L1 and L2 Complex Addition tasks. As shown in Fig. 1a, when hearing the First Number in L2, several cognitive processes take place, which include hearing the L2 word (First Number), rehearsing the L2 number, accessing numerical concepts, and retrieving the L1 translation equivalent. Upon hearing the Second Number in L2, translation, calculation, and working memory related processes are involved. Comparing cognitive components in L1 and L2 Complex Addition, translation is required for solving L2 problems, but not for L1. Other cognitive components are commonly involved in both the L1 and L2 Complex Addition, but L2 tasks could potentially place larger demands on these cognitive components, such as working memory. This is because working memory capacity has been found to be low in bilinguals' L2 and working memory processes are more efficient in L1 (Ardila, 2003).

To further understand L2 mental calculation, we compared 2-digit plus 1-digit addition (*Complex Addition*) and 1-digit plus 1-digit addition (*Simple Addition*, e.g., 2 + 9 = 11?) (Fig. 2). Comparing cognitive components involved during Complex Addition and Simple Addition in L2 (Fig. 1b), carrying operations are required for L2 Complex Addition, but not L1 Simple Addition. Among the common cognitive components, L1 Simple Addition could demand less. For example, there might be minimal translation required for single-digit numbers. Recent studies on bilinguals using EEG measures show that L1 translation equivalents, especially for early-acquired and high-frequency words, are automatically and unconsciously activated when seeing or hearing L2 concrete words (Spalek, Hoshino, Wu, Damian, & Thierry, 2014; Thierry & Wu, 2007; Wu & Thierry, 2012). Recent behavioral data show a frequency effect for number words in which high-frequency words (small numbers) were better recalled in L2 (Sumioka, Williams, & Yamada, 2016). Because single-digit numbers are usually acquired early when learning a foreign language and used with high frequency, the processing load of L2 single-digit numbers could be smaller than L2 multi-digit numbers.

2. Material and methods

2.1. Participants

Twenty-two adult subjects participated in the present study, including eleven native Mandarin Chinese speakers (6 males and 5 females; mean age: 22.0 years, ranging from 19 to 25) and eleven native Japanese speakers (11 males; mean age: 26.54 years, ranging from 21 to 60). All subjects gave their written informed consent in accord with the Institutional Review Board (IRB) at the Veterans General Hospital (Taipei, Taiwan), the Tokyo Denki University (Tokyo, Japan), and the University of Washington (Seattle, U.S.A.). All subjects reported being right-handed. All subjects reported no history of neurological diseases and no arithmetic difficulties. None of the subjects were abacus users.

All participants also completed a questionnaire regarding language background. All the Taiwanese participants reported Mandarin Chinese as their L1 and the preferred language for mental arithmetic. All the Japanese participants reported Japanese as their L1 and preferred language for arithmetic. For the Taiwanese group, the mean onset age of L2 acquisition was 11.32 years (SD = ± 2.24). The self-estimated L2 proficiency level was 3.31 (SD = ± 1.52) (for listening, speaking, reading, and writing on a scale from 1 to 7). The mean percent of English use estimated over the past one year was 14.80% (SD = ± 10.23) across different contexts. These values for the Japanese group were 11.91 years (SD = ± 2.70), 2.42 (SD = ± 1.05), and 13.42% (SD = ± 10.96), respectively.

2.2. Stimuli and experimental design

The auditory verbal number word stimuli were audio recorded from three adult native speakers, one adult Mandarin native speaker, one Japanese, and one American English speaker. The auditory verbal number word stimuli were presented binaurally with inserted earphones. Across all the two-digit number words used in our study, the average durations of number were 820.27 ms ± 89.44 ms in L1 (average of Japanese and Mandarin Chinese ± standard deviation) and 909.35 ms ± 111.45 ms ms in L2. The word duration was significantly shorter in L1 than in L2 ($p < 0.001$). For the single-digit number words, the average durations were 539.10 ± 127.54 ms in L1 (average of Japanese and Mandarin Chinese ± standard deviation) and 595.74 ms ± 99.03 ms in L2. The word duration was not significantly different between L1 and L2 ($p = 0.186$).

Addition problems with “-teen” and “-ty” numbers (e.g., nineteen, ninety) were excluded to avoid potential perceptual confusion in L2 listeners and possible confusion induced by the inversed order of the ones and tens in English and Japanese/Mandarin Chinese (Lin, Imada, & Kuhl, 2012). Addition problems with “one” in the unit digit were also excluded to avoid counting rather than addition. Among the tested problems, 50% of the trials were presented with correct sums. In the other 50% of the trials, the proposed sums were incorrect sums. The proposed sums were presented in L1. By doing this, subjects did not need to translate answers into L2 after retrieving arithmetic facts in L1. To inform subjects of the end of a given trial, Answer was followed by a 500-Hz Warning tone, which was presented 1800 ms after the onset of the Answer stimuli and approximately 1001.17 ms on average after the cessation of the Answer stimuli.

Three tasks were tested in the present study (Fig. 2): (1) two-digit plus one-digit addition task in L1, Japanese or Mandarin Chinese depending on the subjects (L1 Complex Addition), (2) two-digit plus one-digit addition task in L2, English (L2 Complex Addition), and (3) one-digit plus one-digit addition task in L2, English (L2 Simple Addition). Subjects were tested in one language on a given day to avoid confusion and to minimize confounds related to the switching of language modes (Grosjean, 1998; Marian & Spivey, 2003). Tasks and languages were tested in separate sessions in randomized order. In each task, the First Number, the Second Number, and the Answer were presented auditorally and sequentially to find out whether the brain activity differences are observed during the First Number interval and/or during the Second Number interval. Participants were required to respond whether the proposed Answer agreed with the sum of the First and Second Numbers with a response key.

2.3. MEG measurement

MEG measures were collected using a whole-head MEG system (VectorView, Elekta-Neuromag Oy, Finland) for the Taiwanese group and a whole-head MEG system (Neuromag-122, Elekta-Neuromag Oy, Finland) for the Japanese group inside a magnetically shielded room at each site. The VectorView system has 306 sensors with two orthogonal planar gradiometers and one magnetometer at each of the 102 locations, and the Neuromag-122 system houses 122 sensors with two orthogonal planar gradiometers at each of the 61 locations. MEG signals were bandpass filtered between 0.03 and 100 Hz and sampled at 497 Hz. Electrooculogram (EOG) was also recorded simultaneously to detect eye blinks and eye movements.

Subject performance was monitored with online averages during MEG data acquisition. For subjects who tended to have more trials rejected due to high EOG amplitudes, more trials were collected. Thus, the number of trials collected varied across subjects. On average, 170 ± 33 problems were presented for Complex Addition and 189 ± 35 problems for Simple Addition.

2.4. Structural MRI

Structural magnetic resonance images (MRIs) were acquired, using T1-weighted scans, from each Taiwanese subject with a Bruker 3T Medspec300 system (Bruker, Germany), and from each Japanese subject with either a GE 1.5T Signa Excite (GE Medical Systems, USA) or a Hitachi 1.5T Stratis II (Hitachi Medico, Japan).

2.5. Behavioral data analysis

Behavioral responses were collected simultaneously with the MEG measurement. Behavioral performance data (percent accuracy) were compared with paired *t*-tests using two contrasts. To test the effect of language in mental addition, performance in L2 Complex Addition was compared to that in L1 Complex Addition. To test the effect of using 1-digit versus 2-digit numbers on behavioral performance, performance in L2 Complex Addition was compared to that in L2 Simple Addition. Responses to the correct and incorrect proposed answers were collapsed, as the main interest was to examine the behavioral performance differences between L1 and L2. The false discovery rate (FDR) procedure (Benjamini & Hochberg, 1995) was used to control for multiple comparisons.

2.6. MEG data analysis

2.6.1. MEG data preprocessing

Epochs or trials were rejected for further behavioral and MEG analyses if they contained EOG amplitude larger than 450 μV due to blinks or eye movements, or high MEG amplitude larger than 3000 fT/cm (gradiometer) or 3000 fT (magnetometer) due to ambient noises. Trials were rejected if participants made wrong behavioral responses to the Answer or behavioral responses were made later than the Warning tone. Wrong behavioral responses included, for example, pressing the 'Correct' key while an incorrect Answer (e.g., 3 + 4 = 8?) was provided, or pressing the 'Incorrect' key when a correct Answer (e.g., 3 + 4 = 7?) was provided. Trials with no responses or responses made longer after the onset of the Warning tones were excluded from further analysis. This resulted in at least 60 accepted trials for every condition and participant. The recorded MEG responses time-locked to the presentation of the First Number were averaged separately for each task. The averaged data were digitally lowpass filtered at 40 Hz and its DC-offset during the baseline period (-100 to 0 ms) was removed.

2.6.2. MEG source estimates

To obtain cortical activities, each subject's cortical surface was reconstructed using the FreeSurfer software (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999) from the individual MRIs. The forward solutions were calculated based on a realistic boundary element model (BEM) (Hämäläinen & Sarvas, 1989) with a single compartment obtained from each subject's cortical surface. The surface-based source space was defined by the interface between gray and white matter using a recursively subdivided icosahedron by a factor of 5, resulting in about 10,242 vertices or cortical source points per hemisphere.

Using averaged MEG responses, cortical sources were individually estimated at every sampling point from −100 ms before the onset of the First Number to the end of the trial. We used the standardized Low Resolution Brain Electromagnetic Tomography Analysis (sLORETA) inverse method (Pascual-Marqui, 2002), implemented in the Minimum Norm Estimate (MNE) suite software (Gramfort et al., 2014; Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993), without orientation constraints. Noise covariance was estimated from a 100-ms pre-stimulus baseline of all accepted trials. sLORETA allows us to identify distributed patterns of activity during cognitive tasks. Source estimates calculated for each individual subject were then registered and morphed to a study-specific average cortical surface, which was created using the MRIs from all twenty-two subjects who participated in the study.

The magnitude of neural responses has been used to provide evidence for the involvement of a given cortical area for a particular cognitive function (Just, Carpenter, & Miyake, 2003; Kok, 1997). In the present study, the vector norm was first calculated from three-dimensional amplitude values obtained from sLORETA estimates at every cortical source location as a function of time. Then, we defined magnitude measures by normalizing these vector norms with the mean and standard deviation in the baseline period.

To identify the latency ranges that show significant translation and/or calculation effects, brain activity differences between conditions were compared from the onset of the First Number to the end of a given trial. We tested the effect of language by contrasting L2 Complex Addition and L1 Complex Addition. We examined the effect of the 2-digit versus 1-digit addition by contrasting L2 Complex Addition and L2 Simple Addition.

2.6.3. Regions of interest

Time courses of the brain magnitude measures at all source points in ten structural regions of interests (sROIs) were submitted to statistical analysis. These sROIs were pre-determined based on previous neuroimaging studies of arithmetic (Arsalidou & Taylor, 2011; Dehaene et al., 2004; Gruber, Indefrey, Steinmetz, & Kleinschmidt, 2001; Lin et al., 2012; Richard et al., 2000; Salillas, Semenza, Basso, Vecchi, & Siegal, 2012), working memory (Baddeley, 2003; Majerus et al., 2006), translation (Klein et al., 1995; Lehtonen et al., 2005; Rinne et al., 2000), speech (Hickok & Poeppel, 2007), and lexico-semantic processing (Friederici, 2002). These ten sROIs were, bilaterally, the superior temporal sulcus, the pars opercularis of the inferior frontal gyrus, the superior parietal lobule, the intraparietal sulcus, and the supramarginal gyrus. These spatial sROIs were defined anatomically using the Desikan-Killany atlas (Destrieux, Fischl, Dale, & Halgren, 2010) in FreeSurfer.

In our study, trials in the Complex Addition were 5700 ms in time, whereas trials in the Simple Addition condition were 5100 ms in time. To compare conditions of different durations, we equalize durations by using time bins. A constant number of time bins was defined separately in each task so that between-condition comparisons could be performed. For example, in Complex Addition, the First Number interval was rescaled from 1500 ms to 105 time bins. In Simple Addition, the First Number interval was rescaled from 1100 ms to 105 time bins. This resulted in an equal number of time bins (486) for each condition. The average amplitude within a time window was used as the amplitude at a specific time bin. On average, each time bin is about 11.39 ms. When reporting results in time bins in the Results section, we added the corresponding time in ms. Because the neighboring time points were averaged to obtain time bins, the averaging process was similar to applying a temporal smoothing and the temporal details were lost.

2.6.4. Statistical analyses

Due to a large number of statistical tests, cluster-based nonparametric randomization tests were used to correct for multiple comparisons (Maris & Oostenveld, 2007). For the cluster-based nonparametric randomization test, we first obtained the observed t-statistic by comparing the mean magnitude measures across all subjects between tasks within a given sROI at each time point using paired t-tests. Then, all time points passing a primary threshold of $p < 0.001$ were selected. Samples passing the primary threshold were then clustered based on temporal proximity and identified as temporal clusters. In each temporal cluster, t-values were summed. This was followed by the permutation test with 1024 permutations to obtain a null distribution. By the null distribution, permutation p-values were calculated. FDR was used to control for multiple sROIs tested.

2.7. Brain-behavior correlations

To investigate whether performance differences between languages were correlated with brain activity differences, correlations were calculated between performance differences (L1-L2 Complex Addition) and source activity differences (L1-L2 Complex Addition) within the temporal cluster and ROIs that showed significant activity differences. Specifically, we correlated differences of behavioral performance in percent accuracy with differences of brain activity in magnitude. For differences in behavioral performance, we subtracted L2 performance from L1 performance (L1 Complex Addition minus L2 Complex Addition). Differences in brain activity were described by summing magnitude measures both across temporal clusters and across sROIs that showed significant differences in brain activity between the two tasks.

Using the same kind of brain-behavior correlations, we also tested whether performance differences between tasks (L2 Simple Addition minus L2 Complex Addition), which were caused by 2-digit versus 1-digit numbers, co-varied with differences in brain activity. Spearman's rank-order correlation was used in all brain-behavior correlations because it is less sensitive to outliers.

3. Results

3.1. Behavioral performance

In the study, bilingual participants were eleven Japanese-English bilinguals and eleven Taiwanese-English bilinguals. Statistical analyses showed no significant between-group differences were found in the language background questionnaires and the behavioral measures. Neuromagnetic responses in the selected sROIs also showed no significant between-group differences. These two groups of subjects with different L1s confirm that our results are not language specific. The two subject groups were subsequently combined. All twenty-two participants were late unbalanced bilinguals (mean onset age of L2 acquisition: 11.61 years ± 2.44; mean ± SD) with moderate proficiency in their L2 (mean self-estimated L2 proficiency level on a scale from 1 to 7: 2.86 ± 1.36) and the mean percent of English use estimated over the past one year was 14.11% (± 10.37).

Fig. 3 shows the behavioral performance measured in percent accuracy collected simultaneously with the MEG measures. The mean response accuracy from behavioral data was high (88.0%). Subjects made the least errors in the L2 Simple Addition task (92.79% ± 6.79% accuracy; mean ± SD), followed by the L1 Complex Addition task (90.92% ± 11.10%) and the L2 Complex Addition task (80.19% ± 13.61%).

Reaction times were also measured from the onset of the proposed answers to the time point of button presses. The mean reaction time was 1017.092 ± 117.080 ms for L1 Complex Addition, 1090.719 ± 105.272 ms for L2 Complex Addition, 956.0418 ± 112.802 ms for L2 Simple Addition. Reaction time was significantly faster in L1 than L2 Complex Addition, $t(21) = 3.30$, $p < 0.01$ (FDR corrected). Significantly shorter reaction time was also found in L2 Simple Addition compared to L2 Complex Addition, $t(21) = 9.628$, $p < 0.01$ (FDR corrected). However, in our study, subjects were required to respond only after the proposed answers were presented. Thus, the reaction time measured here does not necessarily

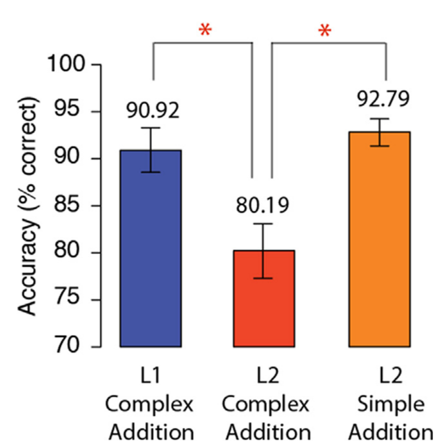

Fig. 3. Behavioral response accuracy obtained during the MEG measurement. Mean percent accuracy and respective standard errors from the three tasks: L1 Complex Addition (blue; Japanese or Mandarin Chinese), L2 Complex Addition (red; English), and L2 Simple Addition (orange; English). Significant differences between languages are marked with asterisks.

reflect the speed of computation. Calculation could have already completed before the presentation of proposed answers.

3.1.1. L2 Complex Addition versus L1 Complex Addition

To examine the behavioral performance differences between L1 and L2, paired *t*-tests were performed on the response accuracy measure. As predicted, significantly higher percent accuracy was found in L1 Complex Addition than L2 Complex Addition, $t(21) = 4.32$, $p < 0.01$ (FDR corrected).

3.1.2. L2 Complex Addition versus L2 Simple Addition

To investigate the role of 2-digit numbers versus 1-digit numbers in behavioral response accuracy, behavioral performance was compared between L2 Complex Addition and L2 Simple Addition tasks. As expected, performance in L2 Complex Addition (80.19%) was significantly lower than in L2 Simple Addition (92.79%), $t(21) = 5.00$, $p < 0.01$ (FDR corrected). L2 Simple Addition (requiring translation) was not significantly different from L1 Complex Addition (90.92%), which required no translation.

3.2. Neuromagnetic responses

3.2.1. L2 Complex Addition versus L1 Complex Addition

We found significantly enhanced activity in L2 Complex Addition compared to L1 Complex Addition in several temporal clusters in the left frontal as well as bilateral parietal areas ($FDR < 0.05$) primarily during the First Number interval (Fig. 4). Specifically, in the left hemisphere sROIs, increased response magnitude was observed in L2 versus L1 in the superior parietal lobule in a temporal cluster at latencies from the 39th to 87th time bin ($p < 0.05$; Cohen's $dz = 0.72$; approximately 457 to 1142 ms), in the pars opercularis of the inferior frontal gyrus in a temporal cluster from the 55th to 110th time bin ($p < 0.05$; $dz = 0.70$; approximately 685 to 1471 ms), and in the supramarginal gyrus in a temporal cluster from the 71st to 108th time bin ($p < 0.05$; $dz = 0.70$; approximately 914 to 1442 ms). In the left intraparietal sulcus, increased activity in L2 was observed in a temporal cluster at an interval from the 48th to 128th time bin ($p < 0.05$; $dz = 0.76$; approximately 585 to 1728 ms), which was a long lasting temporal cluster found during the First Number interval and extending to an early short period of the Second Number interval.

In the right hemisphere sROIs, higher activity levels in the L2 tasks relative to L1 were also observed in the intraparietal sulcus in a temporal cluster from the 43rd to 113th time bin ($p < 0.05$; $dz = 0.73$; approximately 514 to 1514 ms) and in the supramarginal gyrus in a temporal cluster at latency from the 48th to 104th time bin ($p < 0.05$; $dz = 0.76$; approximately 585 to 1385 ms). In the right superior parietal lobule, First Numbers in L2 also evoked greater responses than in L1 in a temporal cluster with latency at the 49th to 119th time bin ($p < 0.05$; $dz = 0.70$; approximately 600 to 1600 ms), which was a large and continuous temporal cluster found during the First Number interval extending to a relatively short and early period of the Second number interval. The long lasting temporal clusters found in the left intraparietal sulcus and right superior parietal lobule during the First Number interval extended to an early short period of the Second Number interval, which was too early for either translation of the Second Number or calculation to take place.

3.2.2. L2 Complex Addition versus L2 Simple Addition

Within sROIs, significantly decreased activity was observed for L2 Simple Addition vs. L2 Complex Addition in several temporal clusters, during the First Number interval ($FDR < 0.05$) (Fig. 5). Some of the continuous temporal clusters found during the First Number interval extended to a relatively brief and early period of the Second number interval, which was too early to be related to calculation. Specifically, we observed increased activity during the First Number interval in the left pars opercularis of the inferior frontal gyrus from the 67th to 113th time bin ($p < 0.05$; $dz = 0.89$; approximately 663 to 1187 ms), from the 65th to 120th time bin ($p < 0.05$; $dz = 0.76$; approximately 640 to 1266 ms) in the left superior temporal sulcus, from the 43rd to 128th time bin ($p < 0.05$; $dz = 0.86$; approximately 389 to 1358 ms) in the right supramarginal gyrus, from the 69th to 119th time bin ($p < 0.05$; $dz = 0.64$; approximately 685 to 1255 ms) in the right superior temporal sulcus, and from the 70th to 133rd time bin ($p < 0.05$; $dz = 0.69$; approximately 697 to 1414 ms) in the right pars opercularis of the inferior frontal gyrus.

3.3. Brain-behavior correlations

3.3.1. L2 Complex Addition versus L1 Complex Addition

To investigate whether differences in behavioral performance between languages co-varied with differences in brain activity differences, we correlated performance differences in accuracy with brain activity differences in magnitude. We found a significant correlation (Spearman's $rho = 0.53$; $p < 0.05$) between behavioral performance differences (from L2 to L1) and brain activity differences (from L2 to L1) (Fig. 6; left).

3.3.2. L2 Complex Addition versus L2 Simple Addition

A significant positive brain-behavior correlation ($rho = 0.60$; $p < 0.01$) was also obtained between behavioral performance differences (Simple minus Complex) and brain activity differences (Simple minus Complex) (Fig. 6; right).

4. Discussion

We utilized MEG to examine the spatiotemporal dynamics of cortical brain responses during mental calculation in bilinguals. Addition problems were presented sequentially and auditorally in bilinguals' L1 and L2. The use of MEG and the experimental paradigm of presenting numbers sequentially allowed us to observe brain activity as a sequence of cognitive processes that unfold over time during mental addition.

4.1. Behavioral findings

Our behavioral data collected simultaneously with MEG recordings showed a significantly higher percentage of accuracy in the Complex Addition task when problems were presented auditorally in L1 as opposed to L2 (Fig. 3). These behavioral results are consistent with previous studies (Marsh & Maki, 1976; McClain & Huang, 1982; Murphey, 2014; Spelke & Tsivkin, 2001). Moreover, performance differences were also obtained when we compared the L2 Complex Addition task and L2 Simple Addition task (Fig. 3). Higher performance accuracy was obtained in the L2 Simple Addition task when one-digit numbers were involved compared to the L2 Complex Addition task when two-digit numbers were involved. No significant performance difference was observed between L2 Simple Addition that required one-digit number translation and L1 Complex Addition that required no translation. This finding in our behavioral data suggests that the translation of one-digit L2 numbers is negligible. During post-experiment interviews, all participants reported translating the L2 numbers into L1 as soon as they heard the numbers. It is less likely that subjects would maintain first operand in L2 in their working memory, hear the second operand in L2, and carry out calculations in their L2. This is because this strategy puts high demands on working memory. Bilingual studies have reported low working memory capacity in bilinguals' L2 (Ardila, 2003) and longer word lengths for numbers in L2 (i.e., English in our present study) compared to L1 (i.e., Japanese or Mandarin Chinese) (Chincotta & Hoosain, 1995; Chincotta & Underwood, 1996; Ellis & Hennelly, 1980). Based on the language background questionnaire, all participants in our experiment indicated L1 as the language they used to learn simple arithmetic and as their preferred language for mental calculation. Thus, in this tested group of late bilinguals with moderate L2 proficiency levels,

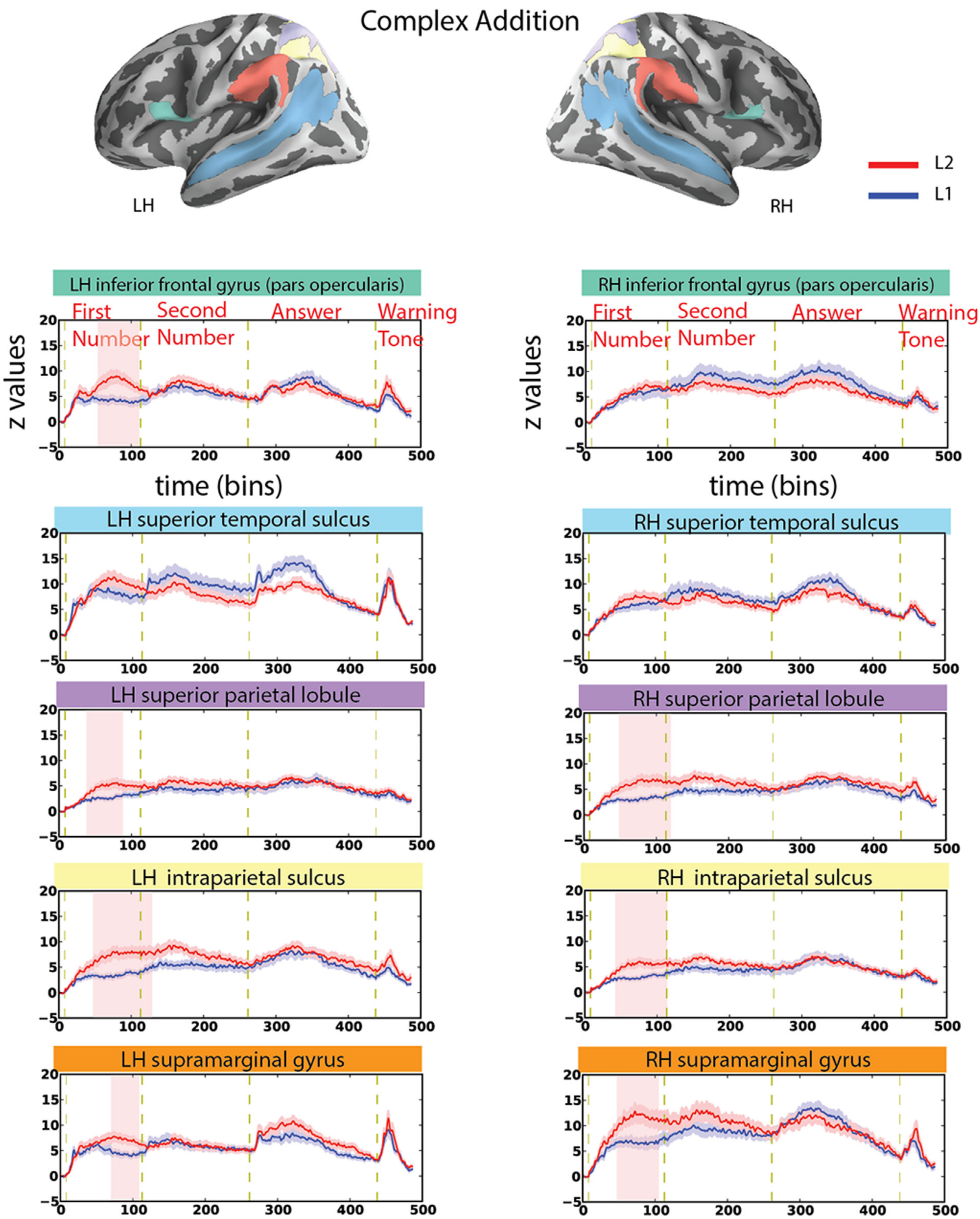

Fig. 4. sROI time courses comparing Complex Addition in L1 versus L2. Measures of mean brain activity magnitude and the standard errors from sROIs are plotted across the whole trial. Selected sROIs are color coded on top of each sROI panel and labeled on the inflated surfaces. Gold dashed vertical lines in each panel mark the beginnings of the First Number, Second Number, Answer, and Warning tone. Significant temporal clusters are marked as a vertical bar in pink.

processing of numbers and arithmetic facts was more accurate when numerical stimuli were presented auditorally in their first language, which is the language they used to acquire basic arithmetic. With early bilinguals who learned both languages simultaneously, we might expect smaller performance differences between languages.

Also, to be noted, arithmetic processing is only one part of mathematical thinking. Performance differences are only expected in tasks that depend on language or verbal components. With tasks not involving verbal components, such as estimation or approximate addition, we would expect relatively small or no language effects (Dehaene et al., 1999; Pica, Lemer, Izard, & Dehaene, 2004). However, verbal codes provide a medium to learn simple arithmetic, and arithmetic is a basic building block of complex and high-order mathematical thinking. Brain activity during basic arithmetic tasks was found to correlate with

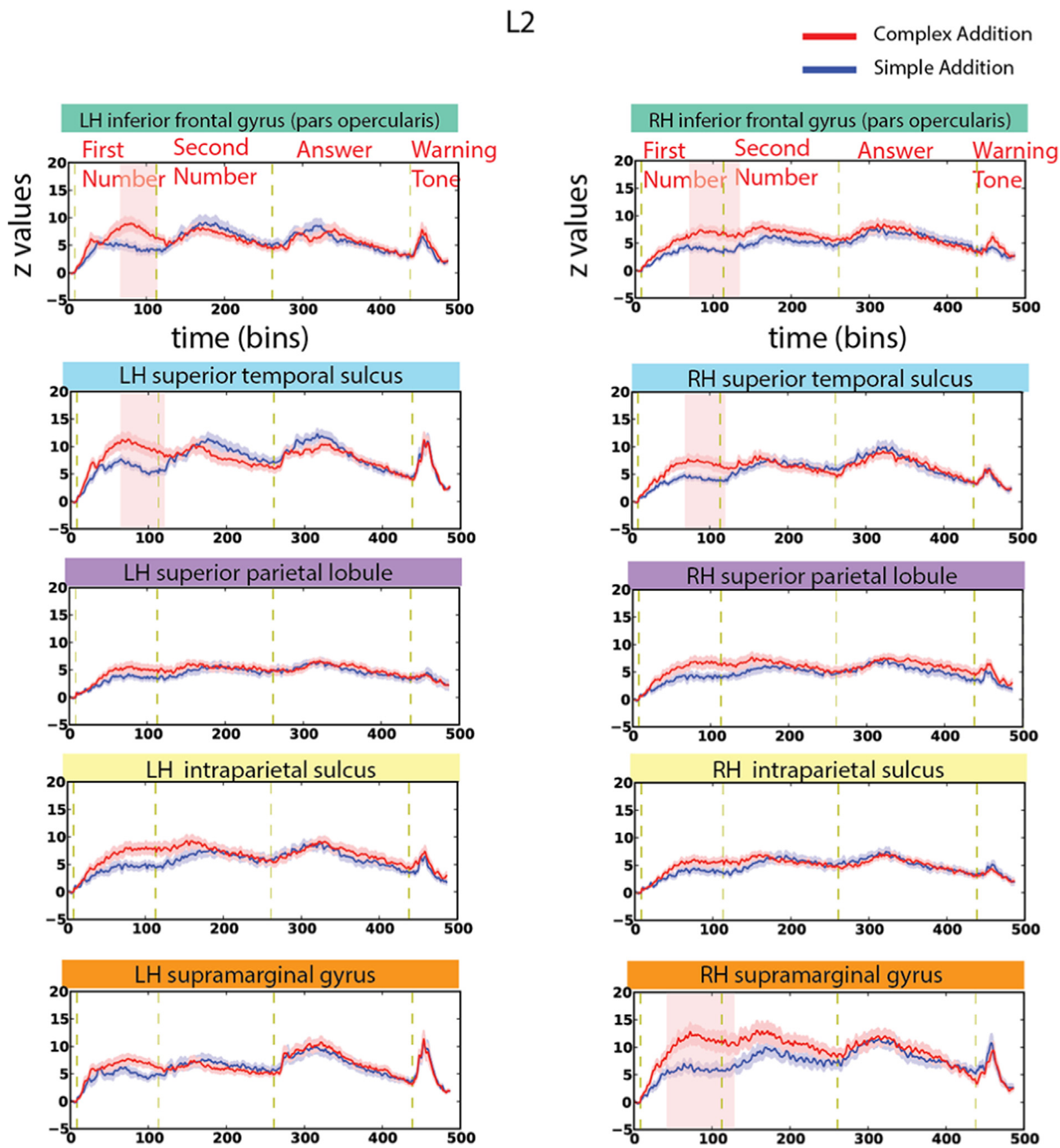

Fig. 5. sROI time courses of sLORETA estimates during the L2 Complex Addition and L2 Simple Addition tasks. Mean sLORETA estimates and the standard errors for the sROIs are plotted across the whole trial, including the First Number, Second Number, Answer, and Warning Tone intervals. See Fig. 3 legend for details of the format.

academic achievement in math (Price, Mazzocco, & Ansari, 2013). Studies on bilingual mental arithmetic thus provide a unique way to probe the role of language in mathematical thinking.

4.2. Neuromagnetic findings

At the cortical level, greater cortical activity was observed in L2 compared to L1 Complex Addition during the First Number interval (Fig. 4). Specifically, enhanced activity was observed in L2 compared to L1 Complex Addition task in inferior frontal areas and bilateral parietal areas. Brain activity differences between L1 and L2 emerged during the First Number interval.

Comparing cognitive components involved, translation is the additional process involved during L2 Complex Addition, but not L1. Increased cortical activity in the left frontal areas during the First Number interval in L2 Complex Addition compared to L1 Complex Addition task could be associated with translation processes. Previous fMRI studies also associated observed inferior frontal activity with the process of translation (Klein et al., 1995; Lehtonen et al., 2005; Rinne et al., 2000), and ERP studies also related slow and sustained activity to translation (Phillips et al., 2006; Thierry & Wu, 2007). Moreover, L2 Complex Addition could place high demands on the common cognitive components, such as working memory processes. Enhanced frontal activity could be associated with actively maintaining the L2 items. Recent MEG data from developmental studies on both infants and adults suggest the role of inferior frontal areas in the formation of internal motor models and possible rehearsal during the auditory presentation of speech sounds (Kuhl et al., 2014). Rehearsal supports temporal

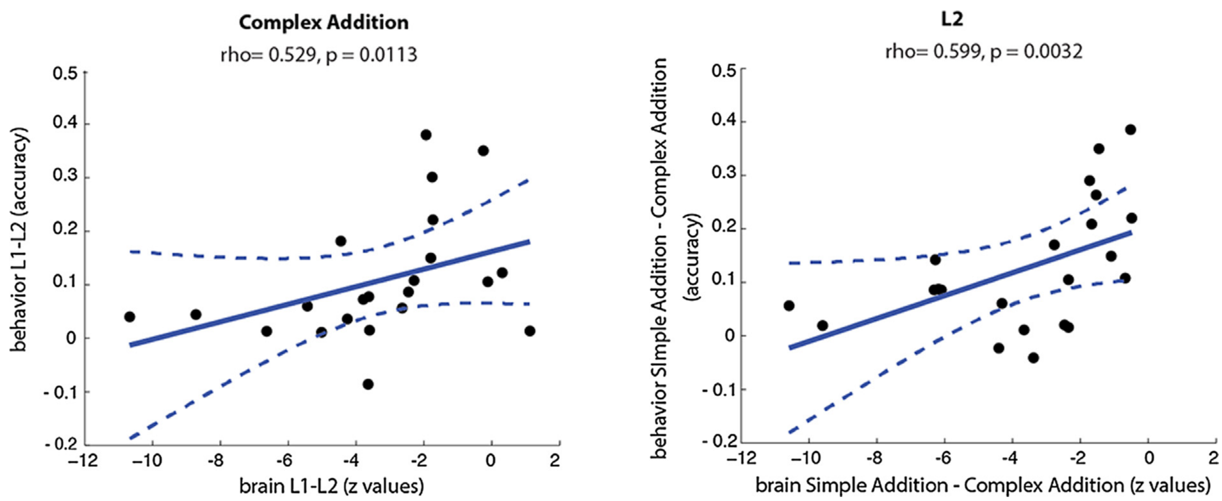

Fig. 6. Brain-behavior correlations between changes in behavioral performance and changes in brain activity. Left: The scatter plot shows a significant positive correlation between behavioral performance difference (L1 minus L2) and brain source activity difference (L1 minus L2). Right: A significant positive correlation is also observed between behavioral performance difference (Simple minus Complex) and source activity difference (Simple minus Complex). The solid line depicts the regression slope and the dashed line depicts the 95% confidence interval. Spearman's rho and associated p values are shown.

retention of L2 words before reaching the L1 translation equivalents.

Increased activity was also found bilaterally in the intraparietal sulcus, superior parietal lobule, and supramarginal gyrus during the First Number interval of L2 Complex Addition compared to L1 Complex Addition task. Increased parietal activity could be associated with increasing demands on retaining L2 numbers, accessing their abstract concepts, and retrieving their translation equivalents. Inferior parietal activity has been shown to reflect numerical quantity retrieval (Eger et al., 2009; Holloway, Price, & Ansari, 2010; Piazza, Pinel, Le Bihan, & Dehaene, 2007). With high spatial resolution, recent studies using high-field fMRI at 7 T (Harvey, Fracasso, Petridou, & Dumoulin, 2015) and ECoG (Daitch et al., 2016) also consistently pointed to the parietal areas as supporting quantity and numerosity processing. In fMRI studies on working memory tasks that involved numerical and order information, left parietal activity has been reported (Attout, Fias, Salmon, & Majerus, 2014; Fegen, Buchsbaum, & D'Esposito, 2015; Majerus et al., 2006). In ERP studies, enhanced activity in later time windows has been associated with tasks involving increased working memory loads (Kiss, Watter, Heisz, & Shedden, 2007; McCollough, Machizawa, & Vogel, 2007; Ruchkin, Johnson, Grafman, Canoune, & Ritter, 1992). Interestingly, previous fMRI studies have reported significant correlations between brain activity in the supramarginal gyrus and verbalization during mental subtraction (Zarnhofer et al., 2012). These results together support our hypothesis that brain activity differences between languages is caused by translation and its related cognitive processes, such as working memory.

During the Second Number interval (Fig. 4), practically no brain activity differences were observed between L2 and L1 Complex Addition tasks. This finding supports our assumption that translation of one-digit L2 numbers is negligible.

By comparing L2 Complex Addition with L2 Simple Addition (Fig. 5), we demonstrated decreased cortical activity during L2 Simple Addition within the First Number interval. Comparing cognitive components involved in these two tasks, enhanced brain activity for L2 Complex Addition could be associated with higher demands on translation and working memory processes. Compared to translating a two-digit number in the Complex Addition task, translating one-digit numbers appears to be highly efficient. One-digit numbers are used with high frequency, are easy to acquire, and learned earlier in the course of learning a new foreign language. Translating one-digit numbers demands less cognitive and neural resources compared to translating two-digit numbers. Moreover, due to the shorter word lengths of one-digit compared to two-digit numbers, working memory loads on maintaining or rehearsing one-digit L2 numbers are also minimal compared to two-digit L2 numbers. We used 2-digit plus 1-digit addition problems for the Simple Addition condition in our study. Future studies should address whether the differences across languages we observed during the First Number interval are minimized or disappear when subjects were tested with 1-digit plus 2-digit addition problems.

4.3. Findings on brain-behavior correlations

Understanding complex processes like mathematical calculations requires studies that relate brain data to behavioral data. The current study produced significant brain-behavior correlations. A significant positive correlation was observed between the L1-L2 performance differences and the L1-L2 brain activity differences during the Complex Addition task (Fig. 6; left). A significant positive brain-behavior correlation was also found between the Simple-Complex performance differences and the Simple-Complex brain activity differences in the L2 task (Fig. 6; right). This result indicates that brain activity differences were associated with behavioral performance differences. As verified by our behavioral and neuromagnetic findings discussed earlier, translation-related activity was negligible during L2 Simple Addition task.

4.4. Perspectives and limitations

Defining bilingualism can be difficult because individuals vary in their proficiency levels, acquisition context, and age of acquisition. Broadly speaking, bilinguals are individuals who use at least two languages (Association, 2004; Birner, 2013). Our participants represent late L2 learners who did not experience math problems in L2 early in life, similar to the situation faced by new immigrant students in schools in the U.S. and Europe. Early bilinguals who learned multiple languages simultaneously may show smaller performance differences between languages, though further studies are needed to determine this.

Several factors should be taken into account when interpreting studies involving mathematical calculations in bilinguals, including the frequency of word usage in L2, the age of onset acquisition of L2, and the proficiency level of L2 (Luk & Bialystok, 2013; Van Rinsveld, Brunner, Landerl, Schiltz, & Ugen, 2015; Van Rinsveld, Dricot, Guillaume, Rossion, & Schiltz, 2017). Studies suggest that both the language used during initial arithmetic learning (Salillas & Wicha, 2012) and current frequency of language use affect brain responses (Martinez-Lincoln et al., 2015). Additional studies are needed to investigate how the frequency of L2 language use impacts behavioral as well as brain activity differences observed between languages.

To be noted, word durations of two-digit numbers in L1 were shorter than those in L2. To maintain the naturalness of spoken number words, we did not artificially equalize durations across number words. The word-length effect has been reported in several studies (Ellis & Hennelly, 1980; Hoosain, 1979; Stigler, Lee, & Stevenson, 1986), showing larger digit spans in the language with shorter word durations. Future studies are required to investigate the effect of word length on bilingual mental calculation. One way to examine the potential word length effect is to test bilinguals with English as L1 and Mandarin Chinese as L2. If similar results are found, for example higher behavioral accuracy in L1 and higher brain activity levels in L2, then we can be assured that the differences are not solely caused by the word length effect.

Our results demonstrate that one-digit number words in L2 induce higher behavioral performance and reduced cortical activity. The Encoding Complex Model (Campbell & Xue, 2001) suggests that bilinguals store arithmetic facts separately for each language and efficiency in accessing a given format or language varies based on prior experience or familiarity. Given L2 one-digit numbers are used frequently, we speculate that frequency of use, and thus practice, potentially helps alleviate difficulties as exhibited by the lower behavioral performance in the two-digit L2 mental addition in bilinguals. More importantly, if frequency of language use matters, further studies should address whether short-term training or within-session practice would minimize performance and activity differences between languages.

4.5. Language experience and mathematical cognition

Our study highlights that experience learning basic arithmetic in a particular language affects arithmetic processing at the behavioral as well as cortical levels. As shown in previous training studies, a switching cost was found when the language used to acquire arithmetic facts differs from the language used for testing. However, human brains are neurally plastic and known to change so as to adapt to the environment and experience. Bilinguals' arithmetic processing can also be modified as experience using another language increases. In monolinguals, shifts in brain activation have been reported as a result of math learning in children and adolescents (Rivera, Reiss, Eckert, & Menon, 2005; Rosenberg-Lee, Barth, & Menon, 2011). In adults, functional changes in brain activity following arithmetic training on complex multiplication problems have also been observed (Grabner et al., 2009; Ischebeck et al., 2006). Bilinguals' experience using a given language is also constantly changing. Thus, as the frequency of language use increases, experience-induced neuroplasticity is expected. Consequently, the switching cost could potentially be reduced as practice with math problems in a second language increases.

5. Conclusions

To conclude, we found that math performance differences between bilinguals' first and second languages arise prior to calculation by using state-of-the-art neuroimaging methods as well as behavioral methods. We demonstrated that L1-L2 brain activity differences were observed during the First Number interval, before the Second Number is presented. We further found that L2 addition problems with 1-digit rather than 2-digit numbers resulted in higher behavioral performance and decreased brain activity. Moreover, we observed that brain activity significantly co-varied with behavioral performance. These results suggest that lower performance when bilinguals calculate in their L2 is associated with a variety of cognitive processes before calculation takes place.

Acknowledgments

This work was supported by a National Science Foundation (NSF) Science of Learning Center grant to the University of Washington's LIFE Center (PKK, PI: SBE-0835854) and by a Washington State Life Sciences Discovery Fund award to PKK. We gratefully acknowledge the contributions of Denise Padden from I-LABS, Chih-Che Chou, Chou-Ming Cheng, and Dr. Jen-Chuen Hsieh from the Taipei Veterans General Hospital, Dr. Keita Tanaka and Dr. Iku Nemoto from the Tokyo Denki University, and all our research participants. We also thank Dr. Juan Silva-Pereyra for valuable feedback on our manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandc.2019.03.006.

References

American Speech-Language-Hearing Association, A. (2004). Knowledge and Skills Needed by Speech-Language Pathologists and Audiologists to Provide Culturally and Linguistically Appropriate Services [Knowledge and Skills]). http://www.asha.org/policy.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders, 36*(3), 233–240.

Arsalidou, M., & Taylor, M. J. (2011). Is 2 + 2 = 4? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage, 54*(3), 2382–2393.

Attout, L., Fias, W., Salmon, E., & Majerus, S. (2014). Common neural substrates for ordinal representation in short-term memory, numerical and alphabetical cognition. *PLoS One, 9*(3), e92049.

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

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57*(1), 289–300.

Bernardo, A. B. (2001). Asymmetric activation of number codes in bilinguals: Further evidence for the encoding complex model of number processing. *Memory & Cognition, 29*(7), 968–976.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859–873.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325–339.

Birner, B. (2013). Bilingualism. *Linguistic Society of America*.

Campbell, J. I., & Epp, L. J. (2004). An encoding-complex approach to numerical cognition in Chinese-English bilinguals. *Canadian Journal of Experimental Psychology, 58*(4), 229–244.

Campbell, J. I., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*(2), 299–315.

Campbell, J. I. D., & Clark, J. M. (1988). An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol, and Goodman (1986). *Journal of Experimental Psychology, 117*, 204–214.

Caramazza, A., & McCloskey, M. (1987). Dissociations of calculation processes. In G. Deloche, & X. Seron (Eds.). *Mathematical disabilities: A cognitive neuropsychological perspective* (pp. 221–234). Hillsdale, NJ: Erlbaum.

Chincotta, D., & Hoosain, R. (1995). Reading rate, articulatory suppression and bilingual digit span. *European Journal of Cognitive Psychology, 7*(2), 201–211.

Chincotta, D., & Underwood, G. (1996). Mother tongue, language of schooling and bilingual digit span. *British Journal of Psychology, 87*(2), 193–208.

Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: Analysis and control in young Swedish-English bilinguals. *Applied Psycholinguistics, 20*, 1–20.

Daitch, A. L., Foster, B. L., Schrouff, J., Rangarajan, V., Kasikci, I., Gattas, S., et al. (2016). Mapping human temporal and parietal neuronal population activity and functional coupling during mathematical cognition. *Proceedings of the National Academy of Sciences, 113*(46), E7277–E7286.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis I. Segmentation and surface reconstruction. *Neuroimage, 9*(2), 179–194.

Dastjerdi, M., Ozker, M., Foster, B. L., Rangarajan, V., & Parvizi, J. (2013). Numerical processing in the human parietal cortex during experimental and natural conditions. *Nature Communications, 4*, 2528.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*(1–2), 1–42.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*, 83–120.

Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology, 14*(2), 218–224.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science, 284*(5416), 970–974.

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage, 53*(1), 1–15.

Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., & Kleinschmidt, A. (2009). Deciphering cortical number coding from human brain activity patterns. *Current*

Biology, 19(19), 1608–1615.
Ellis, N., & Hennelly, R. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71*(1), 43–51.
Estes, K. G., & Hay, J. F. (2015). Flexibility in bilingual infants' word learning. *Child Development, 86*(5), 1371–1385.
Fegen, D., Buchsbaum, B. R., & D'Esposito, M. (2015). The effect of rehearsal rate and memory load on verbal working memory. *Neuroimage, 105*, 120–131.
Fehr, T., Code, C., & Herrmann, M. (2007). Common brain regions underlying different arithmetic operations as revealed by conjunct fMRI-BOLD activation. *Brain Research, 1172*, 93–102.
Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage, 9*(2), 195–207.
Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84.
García-Pentón, L., Garcia, Y. F., Costello, B., Duñabeitia, J. A., & Carrieras, M. (2016). The neuroanatomy of bilingualism: How to turn a hazy view into the full picture. *Language, Cognition and Neuroscience, 31*(3), 303–327.
Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism: Language and Cognition, 6*(1), 1–15.
Gold, B. T., Johnson, N. F., & Powell, D. K. (2013). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia, 51*(13), 2841–2846.
Grabner, R. H., Ansari, D., Reishofer, G., Stern, E., Ebner, F., & Neuper, C. (2007). Individual differences in mathematical competence predict parietal brain activation during mental calculation. *Neuroimage, 38*(2), 346–356.
Grabner, R. H., Ischebeck, A., Reishofer, G., Koschutnig, K., Delazer, M., Ebner, F., et al. (2009). Fact learning in complex arithmetic and figural-spatial tasks: The role of the angular gyrus and its relation to mathematical competence. *Human Brain Mapping, 30*(9), 2936–2952.
Grabner, R. H., Saalbach, H., & Eckstein, D. (2012). Language-switching costs in bilingual mathematics learning. *Mind, Brain, and Education, 6*(3), 147–155.
Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2014). MNE software for processing MEG and EEG data. *Neuroimage, 86*, 446–460.
Grosjean, E. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*, 131–149.
Gruber, O., Indefrey, P., Steinmetz, H., & Kleinschmidt, A. (2001). Dissociating neural correlates of cognitive components in mental calculation. *Cerebral Cortex, 11*(4), 350–359.
Hämäläinen, M., Hari, R., Ilmoniemi, R., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography – Theory, instrumentation, and applications to noninvasive studies of signal processing in the human brain. *Reviews of Modern Physics, 65*, 413–497.
Hämäläinen, M. S., & Sarvas, J. (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Transactions on Biomedical Engineering, 36*(2), 165–171.
Harvey, B. M., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences, 112*(44), 13525–13530.
Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish-English bilinguals: An fMRI study. *Neuroimage, 14*(2), 510–520.
Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience, 8*(5), 393–402.
Holloway, I. D., Price, G. R., & Ansari, D. (2010). Common and segregated neural pathways for the processing of symbolic and nonsymbolic numerical magnitude: An fMRI study. *Neuroimage, 49*(1), 1006–1017.
Hoosain, R. (1979). Forward and backward digit span in the languages of the bilingual. *The Journal of Genetic Psychology, 135*, 263–268.
Iguchi, Y., & Hashimoto, I. (2000). Sequential information processing during a mental arithmetic is reflected in the time course of event-related brain potentials. *Clinical Neurophysiology, 111*(2), 204–213.
Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstatter, F., Benke, T., Felber, S., et al. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. *Neuroimage, 30*(4), 1365–1375.
Jasinski, E. C., & Coch, D. (2012). ERPs across arithmetic operations in a delayed answer verification task. *Psychophysiology, 49*(7), 943–958.
Jost, K., Khader, P., Burke, M., Bien, S., & Rosler, F. (2009). Dissociating the solution processes of small, large, and zero multiplications by means of fMRI. *Neuroimage, 46*(1), 308–318.
Jost, K., Khader, P. H., Burke, M., Bien, S., & Rosler, F. (2011). Frontal and parietal contributions to arithmetic fact retrieval: A parametric analysis of the problem-size effect. *Human Brain Mapping, 32*(1), 51–59.
Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics, 4*, 56–88.
Kiss, I., Watter, S., Heisz, J. J., & Shedden, J. M. (2007). Control processes in verbal working memory: An event-related potential study. *Brain Research, 1172*, 67–81.
Klein, D., Milner, B., Zatorre, R. J., Meyer, E., & Evans, A. C. (1995). The neural substrates underlying word generation: A bilingual functional-imaging study. *Proceedings of the National Academy of Sciences, 92*(7), 2899–2903.
Klein, E., Suchan, J., Moeller, K., Karnath, H. O., Knops, A., Wood, G., et al. (2016). Considering structural connectivity in the triple code model of numerical cognition: Differential connectivity for magnitude processing and arithmetic facts. *Brain Structure and Function, 221*(2), 979–995.
Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: A review and synthesis. *Biological Psychology, 45*(1–3), 19–56.
Kovacs, A. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences, 106*(16), 6556–6560.
Kovacs, A. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science, 325*(5940), 611–612.
Kuhl, P. K., Stevenson, J., Corrigan, N. M., van den Bosch, J. J. F., Can, D. D., & Richards, T. (2016). Neuroimaging of the bilingual brain: Structural brain correlates of listening and speaking in a second language. *Brain and Language, 162*, 1–9.
Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J.-F., & Imada, T. (2014). Infants' brain responses to speech suggest Analysis by Synthesis. *Proceedings of the National Academy of Sciences, USA, 111*, 11238–11245.
Lehtonen, M. H., Laine, M., Niemi, J., Thomsen, T., Vorobyev, V. A., & Hugdahl, K. (2005). Brain correlates of sentence translation in Finnish-Norwegian bilinguals. *Neuroreport, 16*(6), 607–610.
Lin, J. F., Imada, T., & Kuhl, P. K. (2012). Mental addition in bilinguals: An FMRI study of task-related and performance-related activation. *Cerebral Cortex, 22*(8), 1851–1861.
Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605–621.
Luk, G., Bialystok, E., Craik, F. I., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *Journal of Neuroscience, 31*(46), 16808–16813.
Majerus, S., Poncelet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., et al. (2006). The left intraparietal sulcus and verbal short-term memory: Focus of attention or serial order? *Neuroimage, 32*(2), 880–891.
Marian, V., & Fausey, C. (2006). Language-dependent memory in bilingual learning. *Applied Cognitive Psychology, 20*(8), 1025–1047.
Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General, 129*(3), 361–368.
Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition, 6*(2), 97–115.
Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190.
Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory & Cognition, 4*(4), 459–464.
Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition, 11*(1), 81–93.
Martinez-Lincoln, A., Cortinas, C., & Wicha, N. Y. (2015). Arithmetic memory networks established in childhood are changed by experience in adulthood. *Neuroscience Letters, 584*, 325–330.
McClain, L., & Huang, J. Y. (1982). Speed of simple arithmetic in bilinguals. *Memory & Cognition, 10*(6), 591–596.
McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex, 43*(1), 77–94.
Mohades, S. G., Van Schuerbeek, P., Rosseel, Y., Van De Craen, P., Luypaert, R., & Baeken, C. (2015). White-matter development is different in bilingual and monolingual children: A longitudinal DTI study. *PLoS One, 10*(2), e0117968.
Moschkovich, J. (2007). Using two languages when learning mathematics. *Educational Studies in Mathematics, 64*(2), 121–144.
Murphey, D. (2014). The Academic Achievement of English Language Learners: Data for the U.S. and each of the states: Child Trends.
Nunez-Pena, M. I., Cortinas, M., & Escera, C. (2006). Problem size effect and processing strategies in mental arithmetic. *Neuroreport, 17*(4), 357–360.
Nunez-Pena, M. I., Gracia-Bafalluy, M., & Tubau, E. (2011). Individual differences in arithmetic skill reflected in event-related brain potentials. *International Journal of Psychophysiology, 80*(2), 143–149.
Nunez-Pena, M. I., Honrubia-Serrano, M. L., & Escera, C. (2005). Problem size effect in additions and subtractions: An event-related potential study. *Neuroscience Letters, 373*(1), 21–25.
Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232–258.
Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Method and Findings in Experimental Clinical Pharmacology, 24*(D), 5–12.
Phillips, N. A., Klein, D., Mercier, J., & de Boysson, C. (2006). ERP measures of auditory word repetition and translation priming in bilinguals. *Brain Research, 1125*(1), 116–131.
Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron, 53*(2), 293–305.
Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*(5695), 499–503.
Pliatsikas, C., Moschopoulou, E., & Saddy, J. D. (2015). The effects of bilingualism on the white matter structure of the brain. *Proceedings of the National Academy of Sciences, 112*(5), 1334–1337.
Prado, J., Mutreja, R., Zhang, H., Mehta, R., Desroches, A. S., Minas, J. E., et al. (2011). Distinct representations of subtraction and multiplication in the neural systems for numerosity and language. *Human Brain Mapping, 32*(11), 1932–1947.
Price, G. R., Mazzocco, M. M., & Ansari, D. (2013). Why mental arithmetic counts: Brain activation during single digit arithmetic predicts high school math scores. *Journal of Neuroscience, 33*(1), 156–163.
Prieto-Corona, B., Rodriguez-Camacho, M., Silva-Pereyra, J., Marosi, E., Fernandez, T., & Guerrero, V. (2010). Event-related potentials findings differ between children and

adults during arithmetic-fact retrieval. *Neuroscience Letters, 468*(3), 220–224.

Richard, T. C., Romero, S. G., Basso, G., Wharton, C., Flitman, S., & Grafman, J. (2000). The calculating brain: An fMRI study. *Neuropsychologia, 38*, 325–335.

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. *Neuroscience Letters, 294*(2), 85–88.

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex, 15*(11), 1779–1790.

Rosenberg-Lee, M., Barth, M., & Menon, V. (2011). What difference does a year of schooling make? Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving. *Neuroimage, 57*(3), 796–808.

Ruchkin, D. S., Johnson, R., Jr., Grafman, J., Canoune, H., & Ritter, W. (1992). Distinctions and similarities among working memory processes: An event-related potential study. *Cognitive Brain Research, 1*(1), 53–66.

Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction, 26*, 36–44.

Salillas, E., Semenza, C., Basso, D., Vecchi, T., & Siegal, M. (2012). Single pulse TMS induced disruption to right and left parietal cortex on addition and multiplication. *Neuroimage, 59*(4), 3159–3165.

Salillas, E., & Wicha, N. Y. (2012). Early learning shapes the memory networks for arithmetic: Evidence from brain potentials in bilinguals. *Psychological Science, 23*(7), 745–755.

Spalek, K., Hoshino, N., Wu, Y. J., Damian, M., & Thierry, G. (2014). Speaking two languages at once: Unconscious native word form access in second language production. *Cognition, 133*(1), 226–231.

Spelke, E. S., & Tsivkin, S. (2001). Language and number: A bilingual training study. *Cognition, 78*(1), 45–88.

Stigler, J. W., Lee, S. Y., & Stevenson, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition, 23*, 1–20.

Stocco, A., & Prat, C. S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language, 137*, 50–61.

Sumioka, N., Williams, A., & Yamada, J. (2016). Number frequency in L1 differentially affects immediate serial recall of numbers in L2 between beginning and intermediate learners. *Journal of Psycholinguistic Research, 45*(6), 1407–1418.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences, 104*(30), 12530–12535.

Vaid, J., & Menon, R. (2000). Correlates of bilinguals' preferred language for mental computations. *Spanish Applied Linguistics, 4*, 325–342.

Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., & Ugen, S. (2015). The relation between language and arithmetic in bilinguals: Insights from different stages of language acquisition. *Frontiers in Psychology, 6*, 265.

Van Rinsveld, A., Dricot, L., Guillaume, M., Rossion, B., & Schiltz, C. (2017). Mental arithmetic in the bilingual brain: Language matters. *Neuropsychologia, 101*, 17–29.

Venkatraman, V., Siong, S. C., Chee, M. W., & Ansari, D. (2006). Effect of language switching on arithmetic: A bilingual FMRI study. *Journal of Cognitive Neuroscience, 18*(1), 64–74.

Wu, Y. J., & Thierry, G. (2012). Unconscious translation during incidental foreign language processing. *Neuroimage, 59*(4), 3468–3473.

Zarnhofer, S., Braunstein, V., Ebner, F., Koschutnig, K., Neuper, C., Reishofer, G., et al. (2012). The Influence of verbalization on the pattern of cortical activation during mental arithmetic. *Behavioral and Brain Functions, 8*, 13.