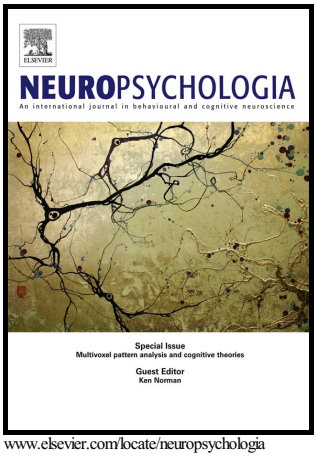
Author's Accepted Manuscript

Behavioral and electrophysiological signatures of word translation processes

Lea B. Jost, Narges Radman, Karin A. Buetler, Jean-Marie Annoni



www.elsevier.com/locate/neuropsychologia

PII: S0028-3932(17)30507-9
DOI: https://doi.org/10.1016/j.neuropsychologia.2017.12.034
Reference: NSY6625

To appear in: *Neuropsychologia*

Received date: 3 July 2017
Revised date: 26 October 2017
Accepted date: 20 December 2017

Cite this article as: Lea B. Jost, Narges Radman, Karin A. Buetler and Jean-Marie Annoni, Behavioral and electrophysiological signatures of word translation processes, *Neuropsychologia,* https://doi.org/10.1016/j.neuropsychologia.2017.12.034

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Behavioral and electrophysiological signatures of word translation processes

Lea B. Jost[a,*], Narges Radman[a], Karin A. Buetler[b], Jean-Marie Annoni[a]

[a]Laboratory for Cognitive and Neurological Sciences, Neurology Unit, Department of Medicine, Faculty of Sciences, University of Fribourg, Fribourg, Switzerland
[b]Leenaards Memory Center, Department of Clinical Neuroscience, Lausanne University Hospital CHUV, Lausanne, Switzerland

Corresponding author*:

Dr. Lea B. Jost
Laboratory for Cognitive and Neurological Sciences
Neurology Unit
Department of Medicine
Faculty of Sciences
University of Fribourg
Ch. du Musée 5
1700 Fribourg

Tel.: +41 26 300 95 85
E-mail address: lea.jost@unifr.ch

Accepted manuscript

Keywords: bilingualism, cognitive control, ERP mapping, translation, word production

Abstract

Translation is a demanding process during which a message is analyzed, translated and communicated from one language to another. Despite numerous studies on translation mechanisms, the electrophysiological processes underlying translation with overt production remain largely unexplored. Here, we investigated how behavioral response patterns and spatial-temporal brain dynamics differ in a translation compared to a control within-language word-generation task. We also investigated how forward and backward translation differs on the behavioral and electrophysiological level. To address these questions, healthy late bilingual subjects performed a translation and a within-language control task while a 128-channel EEG was recorded. Behavioral data showed faster responses for translation

ACCEPTED MANUSCRIPT

compared to within-language word generation and faster responses for backward than forward translation. The ERP-analysis revealed stronger early (<200ms) preparatory and attentional processes for between than within word generation. Later (424-630ms) differences were characterized by distinct engagement of domain-general control networks, namely self-monitoring and lexical access interference. Language asymmetry effects occurred at a later stage (600ms), reflecting differences in conceptual processing characterized by a larger involvement of areas implicated in attention, arousal and awareness for forward versus backward translation.

1. Introduction

As multilingualism plays a crucial role in an increasingly globalized and multicultural world, studies investigating the representation and neural processing of multiple languages have gained substantial interest. A particularly demanding process is translation, where encoding of words (or sentences) takes place in the original language followed by the selective retrieval of the target language.

In a comprehensive meta-analysis, Indefrey and Levelt (2004; updated Indefrey, 2011) identified spatio-temporal correlates for the core processes of within language word generation: accessing the lexical concept of words and lexical selection retrieval around 150-350ms after stimulus onset in anterior middle temporal regions, is followed by posterior middle-temporal phonological code retrieval, posterior inferior-frontal phonological encoding, and articulatory preparation in supplementary motor areas between 350-600ms initiating the articulation of the word. In addition, the model includes self-monitoring processes mainly mediated by superior temporal regions, but also an involvement of a larger network including the cingulate and insular cortex, supplementary and primary motor areas, cerebellum, thalamus and the basal ganglia (Christoffels et al., 2007, van de Ven et al., 2009).

Accepted manuscript

ACCEPTED MANUSCRIPT

In contrast to within language processing, less is known about cross language processing, namely the interaction between the first (L1) and second (L2) language in bilingual word processing. Numerous models of bilingual word access have been brought forward to explain how speakers select words in the target language and how co-activation and thus intrusion from the unintended language can be suppressed. One of the most influential models is the Revised Hierarchical Model (RHM; Kroll & Stewart 1994), which proposes a hierarchical organization of the lexical and the conceptual level, and elaborates the implications of this hierarchical organization for bilingual word production. The RHM distinguishes between two lexicons - one for words of L1 and one for words of L2. These two lexicons are linked to a common conceptual system, which contains the meaning of the words. According to the RHM, L1-L2 translation is slower as compared to L2-L1 translation in bilinguals who acquired their L2 after early childhood and for whom the L1 remains the dominant language. In this model, "both lexical and conceptual links are active in bilingual memory, but the strengths of the links differ as a function of fluency in L2 and relative dominance of L1 to L2" (Kroll & Stewart, 1994, p. 157). As such, the model proposes an asymmetry in the strength of the connections between words and their concepts in the two languages, characterized by stronger links and thus faster access to meaning for L1 words. In bilinguals with a low L2-proficiency, the L2 is assumed to require mediation via the L1 translation equivalent, which in turn leads to slower responses (e.g. Kroll and Stewart 1994; Sholl, Sankaranarayanan and Kroll 1995). The Inhibitory Control Model (IC model, Green 1998), which includes executive-attentional control mechanisms within models of bilingual processing – also referred to as "language control mechanisms" - offers an alternative explanation for this translation-asymmetry. It has been postulated that translation from L1 to L2 requires the inhibition of L1 lemmas in order to produce L2 words. In unbalanced bilinguals, L1 lemmas are assumed to be more active than L2 lemmas, requiring higher attentional resources for the L1 lemmas to be suppressed. Consequently, L1-L2 translation would be slower as compared to L2-L1 translation because the two tasks require differential inhibitory demands (see Sunderman and Kroll 2006; Kroll et al. 2010 for a review). Basic assumptions of the RHM have been

Accepted manuscript

ACCEPTED MANUSCRIPT

called into question (Brysbaert and Duyck 2010) and alternative models of bilingual language selection have been suggested (e.g. Bilingual Interactive Activation BIA+ Model, Dijkstra and van Heuven 2002; Conceptual Selection Model, CSM, Bloem and La Heij 2003), proposing non-selective lexical access across languages. In these models, the presentation of a word leads to automatic activation of lexico-semantic information regarding that word in both languages. The similarity of the input word to the internal lexical representations determines their activation, not the word's language membership. Despite large debate about which linguistic models best explain bilingual word production, there seems to be converging evidence that the parallel activation of two languages gives rise to high demands on cognitive control (e.g. Luk et al. 2012; Abutalebi et al. 2013; Kroll and Bialystok 2013; Kroll et al. 2014).

Together, language models on word processing suggest that early word recognition processes may be similar for first and second language and that the interaction across L1 and L2 may affect higher lexico-semantic, form encoding and attentional processes. La Heij et al (1996) state that differences between L1-L2 and L2-L1 translation "can be accounted for by assuming that to a large extent both tasks are conceptually mediated, but that second-language words are less efficient in activating their concepts than first-language words" (La Heij et al., 1996, p.663). Dijkstra and van Heuven state that "until the target word is identified and its language tag retrieved for responding, cross-linguistic interactions can arise within the mental lexicon" (Dijkstra and can Heuven, 2002, p. 193). In his IC model, Green (1996) emphasizes the role of the "supervisory attentional system", especially in cases where "automatic control is insufficient" (p. 69). Furthermore, the models suggest that the spatio-temporal characteristics of the networks underlying translation production could depend on language proficiency and that unbalanced proficiency may lead to asymmetric bilingual word processing. As Kroll and Biyalistok concluded, "language comprehension and production depend on the absolute levels of proficiency of both languages" (Kroll and Biyalistok, 2013, p. 2). Green further states that "even relatively fluent bilinguals may continue to show an asymmetry in accessing meaning in their two languages" (Green 1998, p. 72).

Accepted manuscript

ACCEPTED MANUSCRIPT

Neuroimaging studies suggest that the involvement of specific neural substrates in translation depends on the source unit (namely words, sentences, supra-sentential text) and the direction of translation (for a review see Garcia 2013; Hervais-Adelman et al. 2011). For single-word translation, studies using spatially sensitive methods have shown inconsistent results. In their PET study, Klein et al. (1995) found that left inferior and dorsolateral frontal and prefrontal regions are activated for both L1-L2 and L2-L1 translation. Similar regions were also activated during synonym generation and rhyme generation tasks. This involvement of the dorsolateral prefrontal cortex was not replicated in the PET-study of Price et al. (1999), who found increased activity in the anterior cingulate, subcortical structures and regions associated with articulation (anterior insula, cerebellum and supplementary motor area) and decreased activation in several other temporal and parietal language areas associated with the meaning of words. According to Price et al. (1999), one reason for these diverging results could be the subjects' different L2-expertise levels, assuming that prefrontal involvement would be stronger when proficiency decreases. Moreover, the studies by Klein et al. (1995) and Price et al. (1999) differed in terms of type of translation (overt vs. silent translation) and baseline task (repetition vs. reading). However, both studies revealed exclusively left hemisphere activation for word translation.

There is also a large number of studies using temporally sensitive methods investigating the electrophysiological processes involved in translation. However, most of these studies used a translation recognition paradigm, in which word pairs are presented in two languages and participants have to indicate whether the second word is the correct translation of the first word (e.g. de Groot 1992; Altarriba and Mathis 1997; Sundermann and Kroll 2006; Guo et al. 2012, Ma et al., 2017). Only few EEG studies have been conducted using an overt translation production task (Janyan et al. 2009; Christoffels et al. 2013). Janyan et al. (2009) investigated the degree of the involvement of semantics in cognate versus non-cognate processing in oral translation of visually presented single words, but for L2-L1 translation only. Christoffels and colleagues (2013) investigated the temporal course of single word translation for both forward (L1-L2) and backward (L2-L1) translation. Their participants were

Accepted manuscript

ACCEPTED MANUSCRIPT

unbalanced but proficient bilinguals, with Dutch as their L1 and English as L2 learned at ~10years of age. Behavioral results showed no reliable differences between translation directions in the reaction times. ERP results indicated that neural processing was different depending on translation direction (L1-L2 or L2-L1) at around 200ms and 400ms following word presentation with larger P2 amplitudes for L1-L2 translation and larger N400 amplitudes for L2-L1 translation. The authors proposed two explanations for this result. As a first option, they suggested that the P2 may be more sensitive to processing the source language whereas the N400 could be associated to lexico-semantic processing related to selecting the target. The authors further postulated that the higher ERP amplitudes, which are often interpreted as reflecting more extensive processing, were associated to higher difficulty to access the L1 input (P2) and produce L1 in translation (N400). They related this last point to the "inversed" language effect also called "paradoxical cost asymmetry" (Meuter and Allport 1999) found in several language-switching studies, which stands for faster responses in L2 than L1 and which has been interpreted in terms of globally reduced access or inhibition of L1 (Costa and Santesteban 2004; Christoffels et al. 2007; Gollan and Ferreira 2009). In the alternative explanation, the authors suggest that the larger P2 amplitude for L1-L2 translation could index the lexical retrieval of the L2 word while the larger N400 for the L2-L1 translation could reflect more effortful lexical and semantic processing of the L2 word. In order to differentiate better between translation processing and reading processes, Christoffels and colleagues (2013) also conducted an experiment for which the material and analyses were based on the translation task, but where participants had to read the words aloud instead of translating them while ERPs were recorded. In contrast to the translation task, no effects of language or interlingual homographs were found. The authors concluded that translation seems to involve conceptual processing in both translation directions and that the task goal influences how words are processed. To our knowledge, the study by Christoffels et al. (2013) is the only one investigating the time course of translation using ERPs and looking at effects of translation direction using an overt production paradigm.

Accepted manuscript

ACCEPTED MANUSCRIPT

The few studies targeting spatio-temporal mechanisms underlying word translation production suggest that the exact nature of the networks may crucially depend on language proficiency. The bigger the asymmetry across L1 and L2, the more brain regions linked to cognitive control may be engaged during word production. In contrast, very little is known about mere spatio-temporal mechanisms underlying translation effects.

The current study takes these previous findings a step further by adding a within-language word generation task, thus providing the possibility to disentangle processes specific for translation from other within-language lexical-semantic processing. Importantly, we chose a word generation task as a control task and not a reading task because in word production, the selection of a lemma is a conceptually driven process, while in reading it is part of the perception process (Indefrey and Levelt 2004). Moreover, the type of ERP analysis used in the present study separates topographic and electrical field strength measures. Differentiating between GFP and topographic effects is meaningful, as different topographies directly indicate different source configurations, whereas different GFP in the absence of a topographic effect indicates different activation strength of the same sources (Michel et al. 2004). This allows a more physiologically oriented interpretation (Murray et al. 2008) of the underlying effects of task and language.

Here, we recorded high-density EEG data during a translation task (L1-L2 and L2-L1) in unbalanced bilinguals with an intermediate proficiency in L2. As a control task, we used a within-language word generation task where participants had to say a word aloud that is semantically related to the presented word. This control task was conducted in both L1 and L2 aiming at disentangling the processes that are specific for translation from within-language processing.

The aim of this study is two-fold:

Aim 1: We investigated which spatial-temporal brain dynamics are different in translation processing compared to within-language processing by comparing behavioral and electrophysiological responses from a translation task to the responses from a within-language word-generation task. On the behavioral level, we expected faster responses for

Accepted manuscript

ACCEPTED MANUSCRIPT

the word-generation task (control task) as compared to the translation task, as no switching from the input to the target language is necessary in the control task. On the electrophysiological level and based on picture naming studies, we hypothesized that the translation task leads to larger amplitudes at around 200-300ms as well as around 400-600ms, as these time windows have previously been associated to lexical retrieval and form encoding, respectively (Indefrey and Levelt 2004). However, considering that response times in translation tasks are slower than in picture naming tasks (e.g. Kroll and Stewart 1994; Cheung and Chen 1998; Francis et al. 2003) lexical retrieval and form encoding might be delayed and occur at later stages. In addition, we expected that lexical retrieval is more difficult in the translation task compared to the control task and that it necessitates a higher load of attentional control mechanisms to switch from the input language to the production language.

Aim 2: Based on studies showing differential effects depending on direction of translation, we investigated the influence of translation direction by comparing the responses of forward translation to responses from backward translation. On the behavioral level we expected that translation would be faster and more accurate for L2-L1 translation than for L1-L2 translation, as the participants in this study had an intermediate level of L2-production skills. At the electrophysiological level we expected similar results as found by Christoffels et al. (2007), namely differential neural processing around 200ms and 400ms following word presentation.

Accepted manuscript

ACCEPTED MANUSCRIPT

2. Materials and Methods

2.1 Participants

26 healthy late bilingual adults participated in the study. Their mother tongue was French (L1), and all participants learned English (L2) after the age of seven (mean = 12.5 years, SD = 2.2) with an intermediate to high proficiency according to the bilingual questionnaire and the two English proficiency tests that were performed (see section “Language evaluation”). All subjects were right-handed and had normal or corrected-to-normal vision. None of the participants reported any history of neurological or psychiatric disorders. From an original group of 26, six participants were excluded due to a low number of correct responses in the language tasks that required a production of a word in L2 (<24/50 correct responses) and three participants because of bad EEG data. Moreover, one participant was excluded because he had misunderstood the task and one due to a technical problem. The final group included 15 healthy late bilingual adults (10 women) aged 18-36 years (mean = 23.4 years, SD = 4.6 years). The sample size was thus comparable to the one reported in the study conducted by Christoffels et al. (2013), including 17 participants for their EEG-analyses. Participants gave written, informed consent and were compensated with 20 CHF/hour for their participation. The study protocol was in agreement with the local ethics committee and conforms to the World Medical Association Declaration of Helsinki. The experiment was carried out in the Neurology Unit, Department of Medicine, University of Fribourg.

2.2 Language evaluation

Participants filled out a custom-made questionnaire assessing their language background including L2-immersion and self-evaluation of their L2-skills (Buetler et al. 2014, 2015). Regarding immersion, participants were asked for the age of L2 acquisition and how present L2 is in their daily activities. For the assessment of L2-presence in current daily activities, participants had to indicate how many percent (25%, 50%, 75% or 100%) they use L2 at work or with friends and how much they consume L2 radio/tv and books (see Table 1). In the

Accepted manuscript

ACCEPTED MANUSCRIPT

self-evaluation section, participants had to indicate on a 10cm scale how good would they estimate their L2 reading, speaking and comprehension skills. Moreover, two tasks were performed to evaluate L2-language proficiency: a sub-test from the computer-based DIALANG language diagnosis system (Zhang and Thompson 2004) to estimate L2 receptive vocabulary and the PVLT (Productive Vocabulary Levels Test, Laufer and Nation 1999) to evaluate L2 productive vocabulary. For the DIALANG, the score ranged between 506 and 1000, with a score >900 indicating native or near-native proficiency and a score between 600 and 900 indicating advanced proficiency with a very substantial vocabulary. A score between 400 and 600 indicates good basic vocabulary, but people at this level may have difficulty handling material that is intended for native speakers. The PVLT samples 18 items at five different word-frequency levels, the first level representing the most frequent 2000 words and the four subsequent levels representing low-frequency words at several vocabulary frequency levels: 3000, 5000, University Word List (UWL) and 10'000. For simplicity, the latter were merged to one score reflecting low frequency words. According to Nation and Waring (1997), the majority of the most frequent 2000 words are content words and knowing them allows a good degree of comprehension of a text. More specifically, second language learners with knowledge of the most frequent 2000 words will know around 80% of the running words in a written or spoken text (Nation and Waring, 1997). Low-frequency words cover the remaining 20%. According to their performance in the PVLT, participants that were included for the analyses in the present study had a knowledge of the most frequent 2000 words of at least 70% and knowledge of low frequency words of at least 38%. Table 1 provides details on participants' L2-proficiency skills and their L2-immersion.

Accepted manuscript

ACCEPTED MANUSCRIPT

Table 1

L2-proficiency skills of participants (n=15)

Variable	mean	SD
Self-evaluation (%)		
Speaking	73	13
Comprehension	84	9
Reading	81	11
Writing	70	15
L2 Vocabulary Tests		
DIALANG score (min = 0; max = 1000)	768	141
PVLT high frequency words (%)	89	9
PVLT low frequency words (%)	53	13
PVLT total score (%)	61	12
Immersion		
Age of acquisition (years)	12.93	2.31
Use of L2 at work/studies (%)	22	18
Consumption of L2 - TV/radio (%)	35	25
Use of L2 with friends (%)	18	14
Consumption of L2-books (%)	35	20

Accepted manuscript

2.3 Tasks and procedure

The participants performed two translation tasks and two within-language word-generation tasks. In the forward translation task, they had to translate L1-words to L2, in the backward translation task they had to translate L2-words to L1. In the within-language word-generation task, words were presented in L1/L2 and the participants were instructed to find a semantically related word in the same language.

Participants were seated in an electrically shielded and sound attenuated booth 120 cm in front of a 22-inch LED screen. Stimuli were presented in the center of the screen and displayed in black font color on light-grey background. Each trial started with the presentation of a fixation cross of 200 ms duration, followed by the randomly determined stimulus displayed for 2000 ms and a 2000-3000 ms interstimulus interval (ISI). The different stimulus conditions (L1-L2 translation, L2-L1 translation, control task L1, control task L2) were presented in blocks. Block order was counterbalanced across subjects. Despite the exclusion of 11 participants in the final analyses, the balancing of block order was still satisfactory. The 15 participants completed the tasks in 14 different block orders. The word-stimuli were all

ACCEPTED MANUSCRIPT

composed of 4-6 letters. In each task, 50 words were presented. Lists of word-stimuli were matched (see Table 2) across and between languages for word type (all nouns), number of letters, log-transformed lexical frequency, neighborhood size, and number of syllables (WordGen, Duyck et al. 2004). All Mean Values as well as p-values are displayed in Table 2. Note that the information given here refers to the language in which the stimuli are presented (i.e. the input language).

At the beginning of each task, a short training block (with four words that were not included in the experimental phase) was started to familiarize the procedure and verify the apparatus, before initiating the experimental phase.

Stimulus delivery and response recording were controlled using E-Prime 2.0 (Psychology Tools, Inc., Pittsburgh, PA, USA). Word productions were recorded and response latencies (time separating the onset of the word presented on the screen and the articulation onset) were systematically checked with a speech analysis software (Audacity®).

Table 2
Mean scores and standard deviations (SD) of the word-stimuli presented in the four tasks, including p-values of the comparisons between the lists.

	Number of letters	Number of syllables	Log-transformed lexical frequency	Neighborhood size
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
List A (words presented in L1 in the translation task)	5.02 (0.73)	1.38 (0.49)	1.62 (0.57)	4.24 (2.88)
List B (words presented in L2 in the translation task)	5.14 (0.75)	1.36 (0.48)	1.49 (0.52)	4.26 (3.97)
List C (words presented in L1 in the control task)	5.02 (0.73)	1.38 (0.49)	1.57 (0.57)	4.28 (2.67)
List D (words presented in L2 in the control task)	5.14 (0.75)	1.40 (0.49)	1.47 (0.49)	4.44 (3.80)
	p-value	**p-value**	**p-value**	**p-value**
List A vs List C	1.000	1.000	0.621	0.943
List B vs List D	1.000	0.684	0.830	0.820
List A vs List B	0.425	0.838	0.226	0.977
List C vs List D	0.425	0.840	0.363	0.809

Accepted manuscript

ACCEPTED MANUSCRIPT

2.4 Electrophysiological recording and data preprocessing

A 128-channel EEG (Biosemi ActiveTwo system, Amsterdam, Netherlands) was recorded at a sampling rate of 1024 Hz. EEG data preprocessing was conducted using Vision Analyzer software (Brain Products GmbH). The data were digitally filtered (lowpass: 40Hz, highpass: 0.1Hz, Notch 50Hz), bad channels were interpolated and eye blinks were corrected using an independent component analysis (Jung et al. 2000). On average, 7.5 (5.9%) electrodes (range=4-12 electrodes) were interpolated. Importantly, for each subject, the same channels were interpolated across all conditions. Trials for which participants gave correct responses and for which RTs were >700ms (to avoid motor artifacts) were epoched 100ms prior and 700ms following the stimulus onset. Trials with artifacts exceeding +-80µV maximal or minimal value within a trial in any channel were automatically rejected. On average, 22% (range 0-50%) of trials were rejected due to incorrect responses or responses faster than 700ms and 7.2% (range 0-28%) were rejected due to artifacts exceeding +-80µV maximal or minimal value within a trial in any channel. Averaging was done separately for each task (translation L1-L2, translation L2-L1, control task L1, control task L2). Only participants whose accepted trial number was equal or above 24 trials in each of the four conditions were included in the analysis. We aimed at having at least 50% (i.e. 25) of the trials accepted, but with a threshold of 24 trials we could increase the group size with 3 additional participants.

2.5 Statistical analyses

2.5.1 Behavior

The average accuracy and response times (RTs) were analyzed using 2x2 repeated measure ANOVAs with the within-subject factors *Task* (translation vs control task) and *Language* (L1 vs L2). Only those trials with response times > 700 ms were included in the analyses in order to make the behavioral data better comparable to the ERP analyses. 12 out of 2328 trials (0.5%) with response times <700 ms were excluded due to possible motor artifacts.

Accepted manuscript

ACCEPTED MANUSCRIPT

Note that the factor Language (L1, L2) refers to the output language (i.e. to the language in which the word was overtly produced). This holds for all analyses reported.

2.5.2 Electrical neuroimaging analysis

For the analysis of the resulting ERPs we used global analyses, namely analyses of global field power (GFP) and topography (Grieder et al., 2012; Buetler et al. 2014; Jost et al. 2014; Sallard et al. 2014). Such analyses have the advantage of being reference-independent (Michel et al. 2004) and of allowing to disentangle if the observed effects are due to differences in strength of activation or due to differences in the configuration of the underlying active brain generators across experimental conditions (e.g. Lehmann 1987; Murray et al. 2008; Michel and Murray 2012; Tzovara et al. 2012).

Differences in strength of the electric field at the scalp were assessed computing a non-parametric randomization test on the GFP (Lehmann and Skrandies 1980; Koenig and Melie-Garcia 2010; Koenig et al. 2011). The GFP represents the spatial standard deviation of the electric field at the scalp and equals the root mean square across all recording electrodes (Lehmann and Skrandies 1980). Potential differences in topography were assessed by applying a point-to-point topographic analyses of variance (TANOVA). TANOVA is based on the global map dissimilarity (GMD) which is calculated as the root mean square of the difference between the strength-normalized voltage potentials across the electrode montage (Lehmann and Skrandies 1980). The TANOVA was performed on the normalized data, which allows attributing topographic differences between conditions to differences in source distribution, and not in source strength (e.g. Grieder et al. 2012). 5000 randomization runs were computed on subject-wise averaged epochs (for details see Koenig and Melie-Garcia 2009; Koenig et al. 2011), with the p-threshold set to 0.05. As multiple time frames were tested, we also computed a test of multiple comparisons in order to know how many consecutive significant time frames can be expected at a chance level of 0.05 (see Grieder et al. 2012). Such statistics on the overall count of significant time points and the duration of the effects are important to prevent false positive results (Koenig et al., 2011). However, early

Accepted manuscript

ACCEPTED MANUSCRIPT

effects are often shorter than later ones and the number of significant time points in a row could thus change over time. As such, when the analysis window is large, they also bare the risk of false negative results, because the number of statistical tests is inflated and over-correction occurs (Koenig & Melie-Garcia, 2010). Nevertheless, in cases with no strong a-priori hypotheses, this rather conservative type of correction for multiple testing can be a valuable tool to prevent over-estimation of results. Both the GFP-analysis as well as the point-to-point TANOVA were applied across all subjects with the within-subject factors *Task* (translation vs control task) and *Language* (L1 vs L2).

In a next step, electrical source estimations were calculated using a distributed linear inverse solution based on a local autoregressive average (LAURA) regularization approach (Grave de Peralta Menendez et al. 2001; Grave-de Peralta et al. 2004; Michel et al. 2004; Sallard et al. 2014). The solution space included 3005 solution points homogenously distributed within the gray matter of the average brain of the Montreal Neurological Institute (courtesy of R. Grave de Peralta Menendez and S. Gonzalez Andino, University Hospital of Geneva, Geneva, Switzerland) and is based on a realistic head model. The results of the above mentioned TANOVA defined the time period over which intracranial sources were estimated. ERPs for each participant and condition (translation L1-L2, translation L2-L1, control task L1, control task L2) were time-averaged over the period showing a significant topographic or GFP difference. Then, intracranial sources were estimated for the resulting one time-sample ERP for each participant and condition and statistically compared at each solution point using a 2x2 ANOVA with the within-subject factors *Task* (translation vs control task) and *Language* (L1 vs L2). Only solutions with a minimal cluster size of 15 points at $p<.05$ were retained in order to control for multiple comparisons (De Lucia et al. 2010; Knebel and Murray 2012). For the resulting regions of interest (ROIs), we extracted and analyzed the T-values, which then allowed drawing conclusions on the direction of the effects, namely whether the solution points found are more or less activated in the conditions of interest.

Accepted manuscript

ACCEPTED MANUSCRIPT

3. Results

3.1. Behavioral Results

ANOVA on behavioral accuracy revealed no difference between the performance in the translation task and the semantic control task (*task* F(1,14) = 0.603, *p*=0.451, partial *η2* = 0.41) but a main effect of *Language* with a higher number of correct responses for L1 as compared to L2 (*language F*(1,14) = 73.545, *p*<0.001, partial *η2* = 0.84). Importantly, this effect was more pronounced for the control task as compared to the translation task (*task x language F*(1,14) = 10.586, *p*<0.01, partial η2 = 0.43). ANOVA on RTs of the correct answers showed faster responses for L1 as compared to L2 (*language F*(1,14) = 95.101, *p*<0.001, partial *η2* = 0.87) and an effect of task, with faster responses in the translation as compared to the control task (*task F*(1,14) = 134.232, *p*<0.001, partial *η2* = 0.90). No interaction was found for RTs (*task x language F*(1,14) = 0.003, *p*<0.955, partial *η*2 = 0.00). See also Figure 1 for illustration.

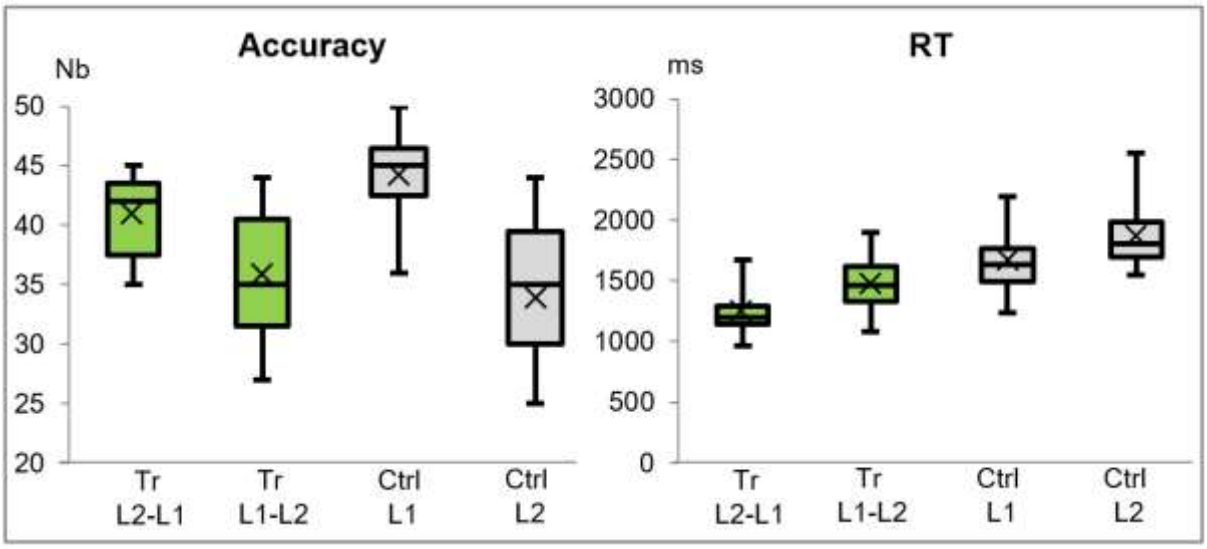


Fig. 1
Behavioral results. Mean accuracy (number of correct responses) and response times (RTs) for the four tasks, namely backward translation (Tr L2-L1), forward translation (Tr L1-L2), control task in L1 (Ctrl L1) and control task in L2 (Ctrl L2).

3.2. Electrophysiological results related to aim 1: comparison between translation and within-language word generation

3.2.1 Event-related potentials

The superimposed ERP waveforms for the four conditions of interest from five exemplary electrodes are depicted in Figure 2.

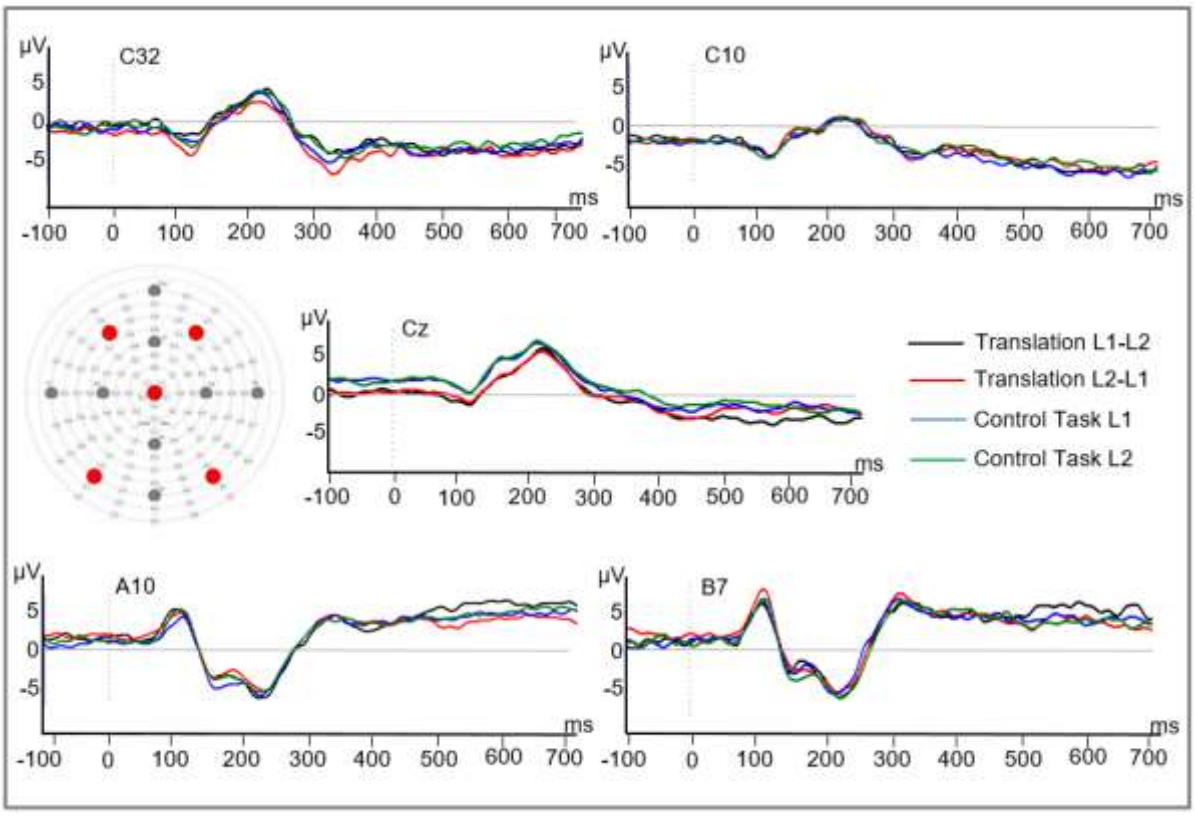


Fig. 2
Superimposed ERP waveforms for the four conditions of interest (translation L1-L2, translation L2-L1, control task L1, control task L2) at five exemplary electrodes.

3.2.2 Global field power

The GFP analysis revealed a main effect of *Task* showing periods of differences ($p<0.05$) between the translation and the control task from -94 to -65ms pre-stimulus onset and 30-83ms post-stimulus onset with higher GFP for the translation compared to the control task. The analysis on the GFP also showed a main effect of *Language* ($p<0.05$) revealing

ACCEPTED MANUSCRIPT

differences between the languages from 165-210ms post-stimulus onset with higher GFP for L2 than L1. However, none of these effects passed the test of multiple comparisons according to which at least 56 consecutive significant time frames can be expected at a chance level of (0.05). Moreover, no significant periods of interaction between the factors *Task* and *Language* were found (all p>0.05). The type of correction for multiple testing is rather conservative. Nevertheless, irrespective of this correction, the time windows found for the main effect of task occurred at the same time as topographic differences, which indicates differences in underlying neural sources (see below). GFP in the presence of topographic differences is less sensitive to measure differences in activation strength originating from specific brain regions. This further indicates that the differences are of topographic nature rather than based on differences in GFP.

3.2.3 Global map dissimilarity

The topographic analysis revealed a main effect of *Task* (p<0.05, >52 TF) identifying significant topographic differences between the translation and the control task from -100-1ms, 24-104ms, 129-203ms and 424-630ms. No main effect of *Language* was found (all p>0.05). The analysis also revealed a significant interaction between *Task* and *Language* from 338-416ms (p<0.05, >52 TF) and from 515-543ms post-stimulus onset (see Figure 3). Additional global duration statistics revealed that 52 consecutive time frames can be expected at the chance level of 0.05, which was the case for the time windows revealing a main effect of *Task* as well as for the time window showing an interaction.

Importantly, the time windows showing a main effect of *Task* do not overlap with the time window showing the interaction. This is noteworthy, as it reveals that the main effect of *Task* is independent of language. In the following, we refrain from further analyzing and interpreting the interaction, as there is a third, nested factor within the factor *Language*, namely "input/output language". However, this nested factor is only present in the translation task, but not in the control task (in which the input and output language are the same). In the

Accepted manuscript

ACCEPTED MANUSCRIPT

following, we thus focus on the main effect of *Task*, which allows us comparing the processes involved in translation with those involved in a within-language word generation task (aim 1). The time windows revealing the significant main effect of *Task* are depicted in Figure 3b.

3.2.4 Electrical source estimations of the main effect of 'Task'

The significant topographic differences between the translation and the control task in the pre-stimulus period of -100-1ms were characterized by stronger activation for the translation task as compared to the control task in the left middle/inferior frontal gyrus (MFG/IFG), in the pre- and postcentral gyrus (PreCG, PostCG), the inferior parietal lobe (IPL) and the superior temporal lobe (STL). Stronger activation for the translation as compared to the control task was also found in the posterior cingulate cortex (PCC), precuneus, cuneus, parahippocampal gyrus (PHG) and lingual gyrus (LG).

Topographic differences in the time period of 24-104ms were characterized by a similar pattern as was the case in the pre-stimulus time range. Additionally, stronger activation for translation compared to the control task was also found in the middle/inferior temporal gyrus (MTG/ITG) and the middle occipital gyrus (MOG).

The topographic differences between the translation and the control task in the time period of 129-203ms were characterized by stronger activation for translation compared to the control task in the left PostCG. In the right hemisphere, the precuneus, the superior and middle temporal gyrus (STG and MTG) and the PreCG were also more activated in the translation as compared to the semantic control task.

Topographic differences in the time period of 424-630ms were characterized by stronger activation for the translation task in the left-hemispheric PreCG and PostCG, the IPL and the STG. Stronger activation for the control task compared to the translation task was found in the medial and superior FG and the anterior cingulate cortex (ACC).

Electrical source estimations are also depicted in Figure 3c.

Accepted manuscript

ACCEPTED MANUSCRIPT

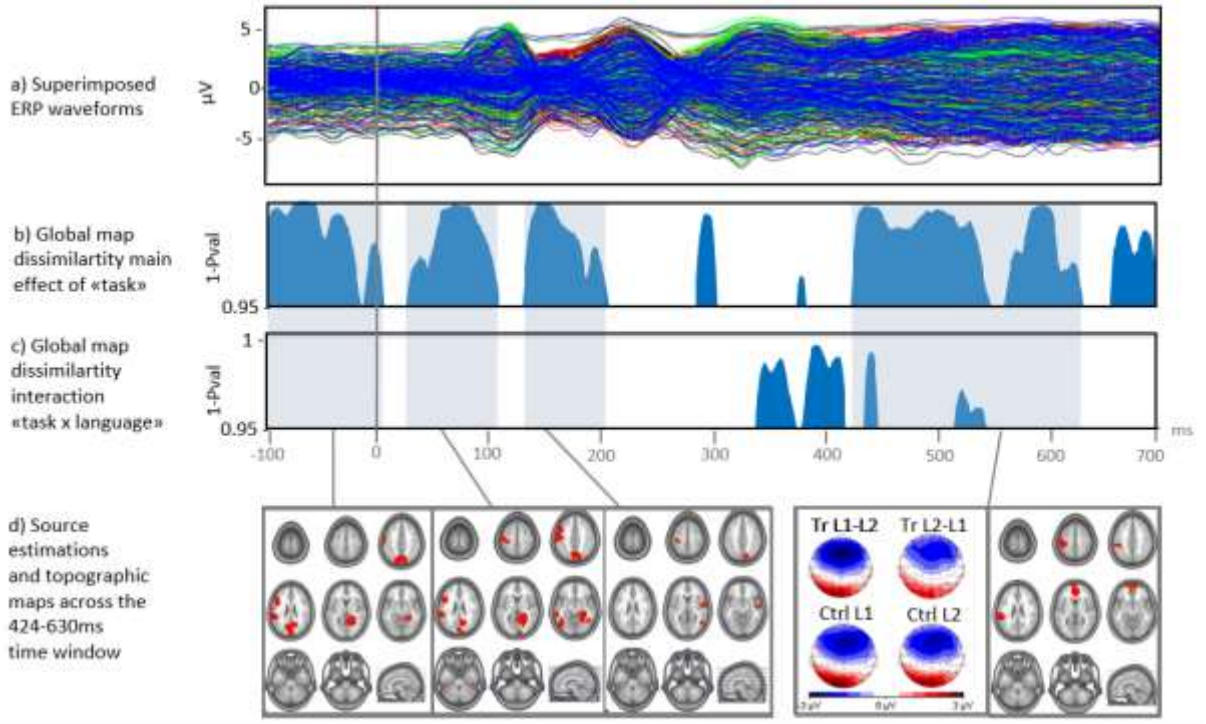


Fig. 3
ERP analyses aim 1. a) Superimposed ERP waveforms for the four conditions of interest across all 128 electrodes. b) Global map dissimilarity of the main effect of t*ask* ($p<0.05$, >52 TF) identifying topographic differences between the translation and the control task from -100-1ms, 24-104ms, 129-203ms and 424-630ms. c) Global map dissimilarity of the interaction *task x language*, revealing significant topographic differences from 338-416ms ($p<.05$, >52 TF). Of note, the time windows revealing a main effect of *task* and the time windows revealing a significant interaction *task x language* do not overlap. d) Source estimations for the four significant time windows revealing a main effect of *task* as well as topographic maps for the four conditions (translation L1-L2, translation L2-L1, control task L1, control task L2) across the 424-630ms window.

3.3 Electrophysiological results related to aim 2: comparison between forward and backward translation

3.3.1 Global field power

The GFP analysis revealed a difference ($p<0.05$) between forward and backward translation from 184-205 ms with higher GFP for L1-L2 translation than for L2-L1 translation. However, this effect did not pass the test of multiple comparisons according to which at least 56 consecutive significant time frames can be expected at a chance level of 0.05.

Accepted manuscript

3.3.2 Global map dissimilarity

The topographic analysis identified significant topographic differences between forward and backward translation from 512-545ms, from 589-618ms and from 647-680ms ($p<0.05$). As the values between the second (589-618ms) and third (647-680ms) time window stayed on a trend-level ($p<0.1$) and the topographic maps were very similar, we merged these two time windows and consider them as one (589-680ms). As additional global duration statistics revealed that 61 consecutive time frames can be expected at the chance level of 0.05, we conducted the electrical source estimation on the second time window only. The time windows showing differences between forward and backward translation are depicted in Figure 4a.

3.3.3 Electrical source estimations

The significant topographic differences between forward and backward translation in the period of 589-680ms were characterized by stronger activation mainly in the posterior cingulate gyrus (PCG) and the thalamus for L1-L2 translation than for L2-L1 translation (see also Figure 4b).

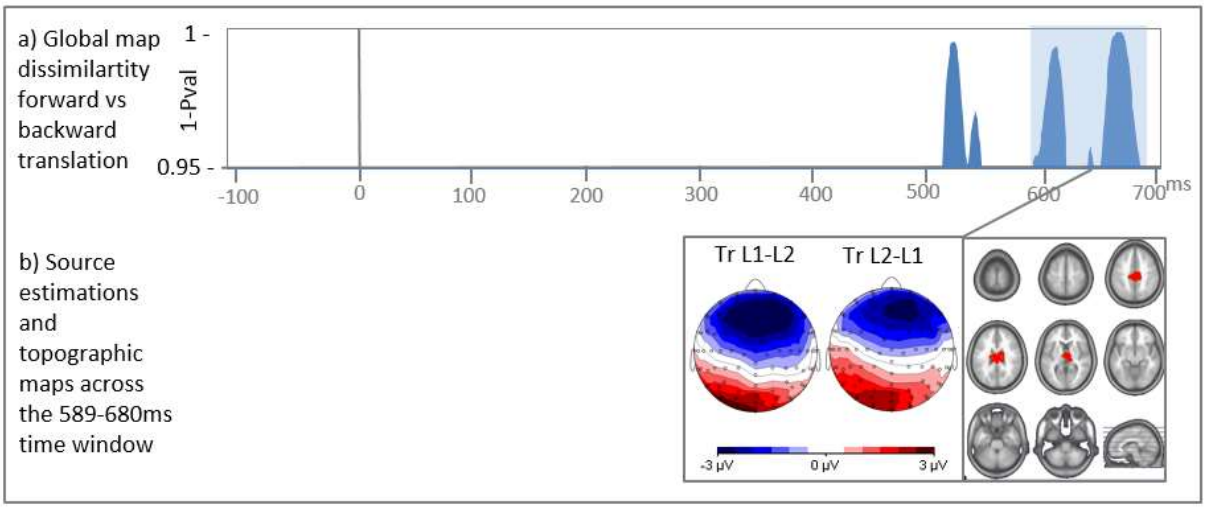


Fig. 4
ERP analyses aim 2. a) Global map dissimilarity identifying significant topographic differences between forward and backward translation ($p<0.05$). The time window passing the global duration statistics (589-680ms, >61 TF) is marked in blue. b) Source estimations for the time window revealing a significant effect of *task* and topographic maps for forward and backward translation across the 589-680ms time window.

ACCEPTED MANUSCRIPT

4. **Discussion**

Translation of words or sentences is a demanding process and has been subject to numerous studies in the area of bilingualism and neuroscience. Yet, only little is known about the electrophysiological processes underlying translation production and in which sense it differs from within-language word generation processing. To our knowledge, this is the first study comparing within-language and between-language word generation using an overt production task and an active control task.

4.1 Aim 1: comparison between translation and within-language word generation

Pursuing aim 1 of the study, we investigated the behavioral and electrophysiological differences between a translation task and a within-language word-generation task in a group of healthy late bilingual participants.

Electrophysiologically, we found differential effects between the translation task and the control task from -100-1ms, 24-104ms, 129-203ms and 424ms-630ms. Importantly, these differential effects were related to differences in topographic distribution of the underlying sources and independent of language.

The differences in the pre-stimulus time range (-100-1ms) most likely reflect voluntary reorientation of attention or response preparation effects, as the different experimental conditions were presented in blocks and therefore predictable for the subjects (Michel et al. 2009). This in turn could have led to the stronger involvement of the largely distributed network including frontal, parietal and temporal areas associated to language and attentional processing during the translation task compared to the within-language word-generation task. In this case, the following time periods (24-104ms and 129-203ms) are likely to be

Accepted manuscript

ACCEPTED MANUSCRIPT

affected by the amount of preparatory potentials, as has previously been shown (Michel et al. 2009), while the late ERP period (424-630ms) stays unaffected by preparatory potentials.

Based on an extensive meta-analysis conducted by Indefrey and Levelt (2004), this late time window (424-630ms) could be interpreted as reflecting differences in form encoding (accessing phonological code, syllabification and phonetic coding) between translation and the semantic control task. In their study, Indefrey and Levelt (2004) proposed a time course of activation for the key-processes of word production. Of note, their conclusions are based on studies using picture naming paradigms in only one language, not considering bilingual language processing. Consequently, the time course suggested by them can only be used as a general guide. Nevertheless, it is likely that the translation and control task involve similar core processes – though probably with longer latency – of word generation. The authors suggest that word production starts by accessing the lexical concept from a visual object around 175ms after picture presentation, followed by lexical selection/lemma retrieval more or less at 150-350ms. Accessing the phonological code, syllabification and phonetic encoding take place between 350-600ms, followed by the articulation of the word. As such, the time period showing differences between the translation and the control task between 424-630ms likely reflect differences in form encoding. More precisely, the retrieval of the corresponding phonological code may be more complex in the translation task involving two languages compared to the within-language task possibly enhancing self-monitoring processes. This interpretation is supported by the analyses performed in the brain space over the same time range (424- 630ms), showing higher activation for the translation than control task in the left PreCG and PostCG, the left IPL and the left STL. The IPL has been associated to bilingual language control, including switching and inhibitory functions (Abutalebi and Green 2007). The left STG seems to play a role in lexical phonological code retrieval around 200-400ms as well as internal self-monitoring after 300ms (Salmelin et al. 1994; Levelt et al. 1998; for a review see Indefrey and Levelt 2004). According to Indefrey and Levelt (2004), the process of self-monitoring is an important feature of speaking, and internal or overt speech seems to act as natural distractors. Thus, the increased difficulty in

Accepted manuscript

ACCEPTED MANUSCRIPT

cross-language phonological code retrieval due to the involvement of two versus one language in the within-language task may enhance internal self-monitoring.

In the same time-range (424- 630ms), the estimated source localization revealed stronger activation for the control task compared to the translation task in the medial and superior frontal gyrus and the anterior cingulate. Previous studies found activation in frontal areas and the ACC for word generation tasks but not picture naming tasks (Indefrey and Levelt 2004). On the other hand, studies with bilingual participants reported ACC-activation also in picture naming tasks, especially when participants had to switch between two languages, likely reflecting increased competition for response selection (Guo et al., 2011, Abutalebi and Green, 2008). For the present study, it thus seems conceivable that the activation of the ACC and frontal areas during word generation reflect stronger competition between co-activated words in the control task than in the translation task. Increased ACC activation is also associated with higher cognitive functions related to cognitive control such as e.g. attention, error and performance monitoring, goal-directed behavior, working memory and inhibition (for a review see Gasquoine 2013), it could be that the control task required more inhibition of the larger amount of words that are automatically activated. A comparable effect was also shown in studies that used a word generation task with different levels of selection demands (e.g. Thompson-Schill et al. 1997; Barch et al. 2000). In these studies, nouns were presented and subjects were asked to produce a verb which is associated to this noun. Results showed that nouns that elicit a larger number of responses activate more strongly the ACC (Barch et al. 2000) and the IFG (Thompson-Schill et al. 1997) as compared to nouns that tend to activate only one verb across subjects. This was interpreted as lower competition in selection/production pathways when only one response is likely to be correct.

The involvement of such language control mechanisms is further supported by our behavioral results showing 1) slower responses for the within-language compared to the translation task, and 2) overall prolonged response latencies in both tasks compared to previous picture naming studies (e.g. Indefrey and Levelt 2004; Valente et al. 2016). Thus, the involvement of additional processes related to attention and interfering with the lexical

Accepted manuscript

ACCEPTED MANUSCRIPT

access seem likely. The slower responses for the within-language word generation task compared to the translation task could result from larger semantic ambiguity for the within word generation task than for translation. This interpretation in terms of semantic ambiguity is also in line with a study showing that interlingual homographs lead to overall slower responses in highly proficient bilinguals, but not in monolingual participants (von Studnitz and Green, 2002).

The presentation of a word may automatically activate several semantically related words. Producing and thus choosing only one of them, which in addition has to be a noun, needs more inhibition than the translation task, for which there is only one or a very limited amount of correct responses. Other studies comparing synonym generation and translation found similar results with longer response latencies for synonym generation compared to translation (e.g. Klein et al. 1995). While in a translation task only few responses are possible, the word-generation task most probably leads to an increased competition between simultaneously activated response possibilities involving a larger within-language cognitive control network, as has previously also been shown for comprehension (e.g. “ambiguity resolution”, Musz and Thompson-Schill 2017). Another interpretation for the finding on increased RT in the control than translation tasks is based on a study comparing within and cross-language ambiguity and suggesting that moderately proficient bilinguals may have direct access to meaning from L2 words (Ma et al., 2017; Degani et al. 2016). As the actual lexical knowledge of bilingual speakers with an intermediate level of L2 proficiency is limited, the translation process could be marked by less ambiguity compared to within-language word-generation, thus leading to faster responses in the translation task.

Not only show our results longer response latencies for the control compared to the translation task, but also overall prolonged response latencies in both tasks compared to previous picture naming studies (e.g. Indefrey and Levelt 2004; Valente et al. 2016), suggesting additional interference with lexical access. For the translation task, this additional process could be linked both to word selection in the other language after semantic activation and to control processes involved in switching from the input to the output language. For the

Accepted manuscript

ACCEPTED MANUSCRIPT

within-language word generation task, this additional process could also involve the control system, needing to resolve the problem of semantic ambiguity and several semantic distractors being activated simultaneously and competing with each other. These control processes could explain the longer response latencies for the tasks used in the present study compared to naming tasks, making it difficult to attribute this time range uniquely to one specific characteristic of the word-production process.

4.2. Aim 2: comparison between forward and backward translation

Aim 2 was to investigate the influence of translation direction by comparing the responses of forward translation to responses from backward translation. Behaviorally, we found slower response times for L1-L2 translation as compared to L2-L1 translation. This is in line with several previous studies in which subjects had a medium L2 level (Kroll and Stewart 1994; Sholl et al., 1995; Francis and Gallard 2005; Francis et al. 2014). L2-proficiency seems to play a crucial role as other studies have shown reverse or no differential effects (e.g. de Groot and Proot 1997; Price et al. 1999; Guo et al. 2012; Christoffels et al. 2013), especially when participants had a relatively high L2-proficiency. In light of the Inhibitory Control Model (Green 1998) this would mean that L1-L2 translation is slower and more error prone as compared to L2-L1 translation due to the difficulty to inhibit the production of the more dominant L1 word. Another possibility would be an interpretation in terms of the RHM model (Kroll and Stewart 1994): L1-L2 translation requires mediation via the L1 translation equivalent, which then leads to the slower responses, especially in bilinguals who are not highly proficient, as is the case of the participants in the present study.

On the electrophysiological level, the GFP-analyses revealed similar results to those of Christoffels et al. (2013), despite using a different approach of ERP-analyses. Stronger activation was found during L1-L2 translation compared to L2-L1 translation around 200ms after word presentation, probably reflecting higher cognitive demand for L2 word generation. However, this effect did not pass the test of multiple comparisons, as the effect did not last

Accepted manuscript

ACCEPTED MANUSCRIPT

long enough. Moreover, the present study revealed topographic differences around 600ms after word-presentation in the absence of differences in strength of activation, while Christoffels et al. (2013) found higher amplitudes for L2-L1 translation around 400ms. These late differences could be related to differential conceptual processing between backward and forward translation. Such an effect has previously been suggested in a study using within- and between language repetition priming (Phillips et al. 2006) revealing that conceptual processing is more strongly involved in L2-L1 translation. Different semantic context effects for L2-L1 translation were also found in a study investigating nonverbal context effects in forward and backward translation (La Heij et al 1996), in which the authors concluded that to a large extent both forward and backward translation are likely conceptually mediated, but that L2 words are less efficient in activating their concepts compared to L1 words.

Source estimations showed that the significant topographic differences between forward and backward translation in the period of 589-680ms were characterized by stronger activation in the PCC and the thalamus for L1-L2 translation as compared to L2-L1 translation. Next to numerous other functions, the thalamus has been associated to coordinating cognitive and motoric aspects of language production (Johnsson and Ojemann 2000). More specifically, it was suggested to play a role in semantic retrieval (e.g. Pergola et al., 2013) as well as motor control of speech and memory retrieval processes (for a review see Johnsson and Ojemann 2000). The PCC is considered to play a crucial role in several intrinsic control networks, in arousal and awareness, episodic memory retrieval and in maintaining a vigilant attentional state (e.g. Pearson et al. 2011; for a review see Leech and Sharp 2014; Raichle 2015). In this sense, larger activation in these areas during L1-L2 translation could thus be reflecting higher attentional demand and higher arousal as forward translation is more difficult compared to backward translation, which is also in line with the behavioral results revealing lower accuracy and slower response times for L1-L2 translation. The PCC has also been associated to the default mode network (DMN, Raichle et al. 2001; Fransson et al. 2008), which plays a role in self-related mental tasks and shows higher activation for tasks that are related to higher self-related mental processing (see Buckner et al. 2008 for a review).

Accepted manuscript

ACCEPTED MANUSCRIPT

The spatial accuracy of EEG inverse solutions remains controversial, as they tend to loose accuracy especially for deep sources (Pascual-Marqui 1999). Scalp-recorded potentials are believed to reflect synchronous activation of neurons that are lined up in parallel (Kutas and Dale 1997). The geometric pattern of deep sources, especially of the thalamus, is said to differ from such a parallel alignment, although there is also the notion that a subset of neurons in the thalamus could possibly generate an electric field that can then be recorded at the scalp (Kounios et al., 2001). We therefore accentuate that the source localizations reported should be interpreted cautiously, especially those regarding the stronger activation in the thalamus for L1-L2 translation compared to L2-L1 translation. However, using a high-density EEG montage with 128 channels, applying statistical parametric mapping analyses to the source estimations (De Peralta-Menendez and Gonzalez-Andino 2002; Michel et al. 2004) as well as applying a conservative statistical approach, this limitation should be at least partially solved. Thus, in light of the scarce amount of electrophysiological studies investigating overt translation, the results can give an indication about the differences in the underlying sources of within- and between language word-generation.

To conclude, our results showed prestimulus and early (<200ms) stronger activation in areas associated to language and attentional processing for the translation compared to the control task, probably related to voluntary reorientation of attention and response preparation effects. The later electrophysiological differences (424-630ms) that were characterized by a distinct activation of global control networks, reflect a different involvement of additional processes related to internal self-monitoring and lexical access interference.

In addition, our results confirm the impact of language asymmetry on the underlying neural mechanisms by replicating previous results showing that L1-L2 translation is slower than L2-L1 translation for bilinguals who are not highly proficient in L2. Electrophysiologically, we found differential processing at a late stage (around 600ms), likely reflecting differences in

Accepted manuscript

ACCEPTED MANUSCRIPT

conceptual processing and characterized by a larger involvement of areas implicated in attention, arousal and awareness for L1-L2 translation.

Acknowledgements

This work was supported by the Swiss National Science Foundation (No. 325130_156937).

Accepted manuscript

Literature

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

Abutalebi J, Green DW. 2008. Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and cognitive processes*, *23*(4), 557-582.

Abutalebi J, Della Rosa PA, Gonzaga, AKC, Keim R, Costa A, Perani D. 2013. The role of the left putamen in multilingual language production. Brain and language. 125(3), 307-315.

Altarriba J, Mathi, KM. 1997. Conceptual and lexical development in second language acquisition. Journal of Memory and Language. 36, 550568. oi:10.1006/jmla.1997.2493.

Barch DM, Braver TS, Sabb FW, Noll DC. 2000. Anterior cingulate and

ACCEPTED MANUSCRIPT

the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. Journal of Cognitive Neuroscience. 12(2), 298-309.

Bloem I, La Heij W. 2003. Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. Journal of Memory and Language. 48, 468–488.

Buetler KA, de León Rodríguez D, Laganaro M, Müri R, Spiere, L, Annoni, JM. 2014. Language context modulates reading route: an electrical neuroimaging study. Frontiers in Human Neuroscience. 8, 83, doi: 10.3389/fnhum.2014.00083.

Buetler KA, de León Rodríguez D, Laganaro M, Müri R, Nyffeler T, Spierer L, Annoni JM. 2015. Balanced bilinguals favor lexical processing in their opaque language and conversion system in their shallow language. Brain and language. 150, 166-176.

Brysbaert M, Duyck W. 2010. Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? Bilingualism: Language and cognition. 13(03), 359-371.

Buckner RL, Andrews-Hanna JR, Schacter DL. 2008. The brain's default network. Annals of the New York Academy of Sciences, 1124(1), 1-38.

Cheung H, Chen HC. 1998. Lexical and conceptual processing in Chinese-English bilinguals: Further evidence for asymmetry. Memory & Cognition. 26(5), 1002-1013.

Christoffels IK, Firk C, Schiller N. O. 2007. Bilingual language control: An event-related brain potential study. Brain research. 1147, 192-208.

Christoffels IK, Ganushchak L, Koester D. 2013. Language conflict in

Accepted manuscript

ACCEPTED MANUSCRIPT

translation: an ERP study of translation production. Journal of Cognitive Psychology. 25(5), 646-664.

Costa A, Santesteban M. 2004. Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. Journal of Memory and Language. 50, 491511. doi:10.1016/j.jml.2004.02.002.

Degani T, Prior A, Eddington CM, da Luz Fontes ABA, Tokowicz N. 2016. Determinants of translation ambiguity. Linguistic approaches to bilingualism. 6(3), 290-307.

de Groot AMB. 1992. Determinants of word translation. Journal of Experimental Psychology: Learning, Memory, and Cognition. 18, 10011018. doi:10.1037/0278-7393.18.5.1001.

De Groot AMB, Poot R. 1997.Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. Language Learning. 47, 215–264.

De Lucia M, Clarke S, Murray MM. 2010. A temporal hierarchy for conspecific vocalization discrimination in humans. J. Neurosci. 30, 11210–11221. doi: 10.1523/JNEUROSCI.2239-10.2010.

De Peralta-Menendez RG, Gonzalez-Andino SL. 2002. Comparison of algorithms for the localization of focal sources: evaluation with simulated data and analysis of experimental data. International Journal of Bioelectromagnetism. 4(1).

Dijkstra T, Van Heuven WJ. 2002. The architecture of the bilingual word recognition system: From identification to decision. Bilingualism: Language and cognition. 5(03), 175-197.

Accepted manuscript

ACCEPTED MANUSCRIPT

Duyck W, Desmet T, Verbeke L, Brysbaert M. 2004. WordGen: a tool for word selection and non-word generation in Dutch, German, English, and French. Behav. Res. Methods Instrum. Comput. 36, 488–499. doi: 10.3758/BF03195595.

Francis WS, Augustini BK, Sáenz SP. 2003. Repetition priming in picture naming and translation depends on shared processes and their difficulty: evidence from spanish-english bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition. 29(6), 1283.

Francis WS, Gallard SLK. 2005. Conceptmediation in trilingual translation: Evidence from response time and repetition priming patterns. Psychonomic Bulletin & Review. 12, 1082–1088.

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

Fransson P,Marrelec G. 2008. The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. Neuroimage. 42(3), 1178-1184.

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

Gasquoine PG. 2013. Localization of function in anterior cingulate cortex: from psychosurgery to functional neuroimaging. Neuroscience & Biobehavioral Reviews. 37(3), 340-348.

Gollan TH, Ferreira VS. 2009. Should I stay or should I switch? A cost-benefit

Accepted manuscript

ACCEPTED MANUSCRIPT

analysis of voluntary language switching in young and aging bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition. 35, 640665. doi:10.1037/a0014981.

Grave de Peralta Menendez R, Gonzalez Andino S, Lantz G, Michel CM, Landis T. 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. Brain Topogr. 14, 131–137.

Grave-de Peralta R, Gonzalez-Andino S, Gomez-Gonzalez CM. 2004. The biophysical foundations of the localisation of encephalogramgenerators in the brain. The application of a distribution-type model to the localisation of epileptic foci. Rev. Neurol. 39, 748–756.

Green DW. 1998. Mental control of the bilingual lexicosemantic system. Bilingualism: Language and Cognition. 1, 67–81.

Grieder M, Crinelli RM, T, Wahlund LO, Dierks T, Wirth M. 2012. Electrophysiological and behavioral correlates of stable automatic semantic retrieval in aging. Neuropsychologia. 50(1), 160-171.

Guo T, Liu H, Misra M, Kroll JF. 2011. Local and global inhibition in bilingual word production: fMRI evidence from Chinese–English bilinguals. *NeuroImage*, *56*(4), 2300-2309.

Guo T, Misra M, Tam JW, Kroll JF. 2012. On the time course of accessing meaning in a second language: An electrophysiological and behavioral investigation of translation recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 38, 11651186. doi:10.1037/a0028076

Hervais-Adelman AG, Moser-Mercer B, Golestani N. 2011. Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to. Bilingualism and cognitive control. 29.

Accepted manuscript

ACCEPTED MANUSCRIPT

Indefrey P, Levelt WJ. 2004. The spatial and temporal signatures of word production components. Cognition. 92(1), 101-144.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. Frontiers in psychology, 2, 255.

Janyan A, Popivanov I, Andonova E. 2009. Concreteness effect and word cognate status: ERPs in single word translation. Brain Talk, 21.

Johnson MD, Ojemann GA. 2000. The role of the human thalamus in language and memory: evidence from electrophysiological studies. Brain and cognition. 42(2), 218-230.

Jost LB, Eberhard-Moscicka AK, Frisch C, Dellwo V, Maurer U. 2014. Integration of spoken and written words in beginning readers: A topographic ERP study. Brain topography. 27(6), 786-800.

Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. 2000 Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin Neurophysiol. 111(10):1745–1758. doi:S1388-2457(00)00386-2

Klein D, Milner B, Zatorre RJ, Meyer E, Evans AC. 1995. The neural substrates underlying word generation: a bilingual functional-imaging study. Proceedings of the National Academy of Sciences. 92(7), 2899-2903.

Knebel JF, Murray MM. 2012. Towards a resolution of conflicting models of illusory contour processing in humans. Neuroimage. 59, 2808–2817. doi: 10.1016/j.neuroimage.2011.09.031

Koenig T, Melie-Garcia L. 2010. A method to determine the presence of averaged event-related fields using randomization tests. Brain Topography. 23, 233–242.

Accepted manuscript

ACCEPTED MANUSCRIPT

Koenig T, Melie-Garcia L. 2009. Statistical analysis of multichannel scalp field data. In C. M. Michel, T. Koenig, D. Brandeis, L. R. R. Gianotti, & J. Wackermann (Eds.), Electrical Neuroimaging (pp. 169–190). New York: Cambridge University Press.

Koenig T, Kottlow M, Stein M, Melie-Garcia L. 2011. Ragu: A free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. Computational Intelligence and Neuroscience, 2011, 938925.

Kounios J., Smith RW., Yang W., Bachman P. D'Esposito M. 2001. Cognitive association formation in human memory revealed by spatiotemporal brain imaging. Neuron, 29(1), 297-306.

Kutas M, Dale AM. 1997. Electrical and magnetic readings of mental functions. Cognitive Neuroscience, 197-242.

Kroll JF, Stewart E. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. Journal of memory and language. 33(2), 149.

Kroll JF, Van Hell JG, Tokowicz N, Green DW. 2010. The Revised Hierarchical Model: A critical review and assessment. Bilingualism: Language and Cognition. 13, pp 373-381 doi:10.1017/ S136672891000009X.

Kroll JF, & Bialystok E. 2013. Understanding the consequences of bilingualism for language processing and cognition. Journal of Cognitive Psychology. 25(5), 497-514.

Kroll JF, Bobb SC, Hoshino N. 2014. Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. Current directions in psychological science. 23(3), 159-163.

Accepted manuscript

ACCEPTED MANUSCRIPT

La Heij W, Hooglander A, Kerling R, Van Der Velden E. 1996. Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. Journal of Memory and Language. 35(5), 648-665.

Laufer B,Nation P. 1999. A vocabulary-size test of controlled productive ability. Lang. Test.16 (1),33–51.

Leech R, Sharp DJ. 2014. The role of the posterior cingulate cortex in cognition and disease. Brain. 137(1), 12-32.

Lehmann D. 1987. Principles of spatial analysis. In: Gevins, A.S., Remond, A. (Eds.), Handbook of Electroencephalography and Clinical Neurophysiology. Methods of Analysis of Brain Electrical and Magnetic Signals, vol. 1. Elsevier, Amsterdam, pp. 309–354.

Lehmann D, Skrandies W. 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalography and Clinical Neurophysiology. 48, 609–621.

Levelt WJM, Praamstra P, Meyer AS, Helenius P, Salmelin R. 1998. An MEG study of picture naming. Journal of Cognitive Neuroscience. 10, 553–567.

Luk G, Green DW, Abutalebi J, Grady C. 2012. Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. Language and cognitive processes. 27(10), 1479-1488.

Ma F, Chen P, Guo T, Kroll JF. 2017. When late second language learners access the meaning of L2 words: using ERPs to investigate the role of the L1 translation equivalent. Journal of Neurolinguistics. 41, 50-69.

Meuter RF, Allport A. 1999. Bilingual language switching in naming:

Accepted manuscript

ACCEPTED MANUSCRIPT

Asymmetrical costs of language selection. Journal of memory and language. 40(1), 25-40.

Michel CM, Murray MM, 2012. Towards the utilization of EEG as a brain imaging tool. NeuroImage. 61, 371–385.

Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, de Peralta RG. 2004. EEG source imaging. Clinical neurophysiology. 115(10), 2195-2222.

Michel CM, Koenig T, Brandeis D, Gianotti LRR, Wackermann J.(2009. Electrical Neuroimaging. New York: Cambridge University Press.Michel, C.M., Murray MM, Brunet D, Michel CM. 2008. Topographic ERP analyses: A stepby-step tutorial review. Brain Topography. 20, 249–264.

Musz E, Thompson-Schill SL. 2017. Tracking competition and cognitive control during language comprehension with multi-voxel pattern analysis. Brain and Language. 165, 21-32.

Nation P, Waring R. 1997. Vocabulary size, text coverage and word lists. *Vocabulary: Description, acquisition and pedagogy*, *14*, 6-19.

Pascual-Marqui RD. 1999. Review of methods for solving the EEG inverse problem. International Journal of Bioelectromagnetism. 1(1), 75-86.

Pearson JM, Heilbronner SR, Barack DL, Hayden BY, Platt ML. 2011. Posterior cingulate cortex: adapting behavior to a changing world. Trends in cognitive sciences. 15(4), 143-151.

Pergola G, Bellebaum C, Gehlhaar B, Koch B, Schwarz M, Daum I, Suchan B. 2013. The involvement of the thalamus in semantic retrieval: a clinical group study. *Journal of cognitive neuroscience, 25(6),* 872-886.

Phillips NA, Klein D, Mercier J, de Boysson C. 2006. ERP measures

Accepted manuscript

ACCEPTED MANUSCRIPT

of auditory word repetition and translation priming in bilinguals. Brain Research. 1125(1), 116-131.

Price CJ, Green DW, Von Studnitz R. 1999. A functional imaging study of translation and language switching. Brain. 122(12), 2221-2235.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. Proceedings of the National Academy of Sciences. 98(2), 676-682.

Raichle ME. 2015. The brain's default mode network. Annual review of Neuroscience. 38, 433-447.

Sallard E, Barral J, Chavan CF, Spierer L. 2014. Early attentional processes distinguish selective from global motor inhibitory control: an electrical neuroimaging study. NeuroImage. 87, 183-189.

Salmelin R, Hari R, Lounasmaa OV, Sams M. 1994. Dynamics of brain activation during picture naming. Nature. 368, 463–465.

Sholl A, Sankaranarayanan A, Kroll JF. 1995. Transfer between picture naming and translation: A test of asymmetries in bilingual memory. Psychological Science. 6, 45–49.

Sunderman G, Kroll JF. 2006. First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. Studies in second language acquisition. 28(03), 387-422.

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. Proceedings of the National Academy of Sciences. 94(26), 14792-14797.

Tzovara A, Murray MM, Michel CM, De Lucia M. 2012. A tutorial review of

Accepted manuscript

ACCEPTED MANUSCRIPT

electrical neuroimaging from group-average to single-trial event-related potentials. Dev. Neuropsychol. 37, 518–544.

Valente A, Pinet S, Alario FX, Laganaro M. 2016. “When” does picture naming take longer than word reading? Frontiers in psychology. 7.

van de Ven, V., Esposito, F., & Christoffels, I. K. (2009). Neural network of speech monitoring overlaps with overt speech production and comprehension networks: a sequential spatial and temporal ICA study. Neuroimage, 47(4), 1982-1991.

Von Studnitz RE, Green DW. 2002. The cost of switching language in a semantic categorization task. *Bilingualism: Language and Cognition*, *5*(3), 241-251.

Zhang S, Thompson N. 2004, "DIALANG: A Diagnostic Language Assessment System (review)." The Canadian Modern Language Review. 61 (2): 293.

Accepted manuscript

ACCEPTED MANUSCRIPT

Highlights

Behavioral responses are faster for backward than forward translation

Responses are faster in simultaneous translation than within-language word generation

Neural attentional and self-monitoring effects are stronger in translation

Lexical access interference needs more control in within-language word generation

Late language asymmetry effects reflect differences in conceptual processing

Accepted manuscript