

Ann-Kathrin Habig*

How does interpreting training affect the executive function of switching? A longitudinal EEG-study of task switching

<https://doi.org/10.1515/psicl-2024-0093>

Received November 19, 2024; accepted May 29, 2025

Abstract: How experience-dependent plasticity can lead to structural and functional brain changes has long been examined in cognitive studies, leading neurolinguistics to investigate how brain structure and functionality are affected by training and controlling multiple languages in a similar manner. Overlapping networks of language control and domain-general executive functions are implicated in this process of managing elevated control demands in coordinating more than one language. Simultaneous interpreting (SI), as a cognitively extremely taxing way of controlling languages, holds a special status. SI involves high processing demands and frequent switching between languages under time pressure. Therefore, SI offers important insight on how task switching abilities may change through training language control in a professional setting. Within a longitudinal design using event-related potentials, this study examines how the executive function of switching progresses during interpreting training. Results show a progression of voltage in the P300 time window until Session 4, followed by a decrease that indicates an automation of attention switching and working memory updating. No significant effects were found within the N200 time-window, which only allows for observing a trend towards progression of N200 over time, suggesting the development and subsequent automation of inhibitory control aspects of switching induced by interpreting training.

Keywords: executive functions; experience-dependent plasticity; interpreting; task switching; ERP

1 Introduction

Managing multiple languages has been shown to entail both structural as well as functional changes in several areas of the brain, highlighting the brain as a distinctly

***Corresponding author: Ann-Kathrin Habig**, English Linguistics and Translation Studies, TraCo Lab, Johannes Gutenberg University Mainz, FB06 Germersheim, Mainz, Germany, E-mail: anhabig@uni-mainz.de

adaptive system that can adjust itself flexibly to different experiences of language use. Mechelli et al. (2004) found an increase in grey matter density in the left inferior parietal cortex of early and late bilinguals compared to monolinguals, with a more pronounced increase in early bilinguals compared to late bilinguals. Furthermore, their results also showed that grey matter density in the observed region increased with L2 (second language) proficiency but declined with increasing age of acquisition. Grogan et al. (2012) examined lexical efficiency in multilinguals compared to bilinguals and found increased grey matter density in the right posterior supra-marginal gyrus for multilinguals as well as a negative relation of grey matter volume in this region to age of acquisition. Results from both studies indicate that these areas are related to the number of words that are learned in monolinguals compared to bilinguals and multilinguals. Similarly, Hosoda et al. (2013) showed that grey matter volume in the inferior frontal gyrus pars opercularis increased with increased L2 vocabulary competence in late L2 learners. Furthermore, connectivity of the inferior frontal gyrus pars opercularis with the caudate nucleus and the superior temporal gyrus increased with growing L2 vocabulary competence. Subcortically, Pliatsikas et al. (2017) found reshaping in the right and caudate nucleus, bilateral thalamus as well as putamen that was positively correlated to L2 immersion of sequential bilinguals, while sequential bilinguals with comparable L2 proficiency and age of acquisition that were less immersed or used L2 less frequently did not show these results. Rossi et al. (2017) showed that L2 learning had a significant effect on connectivity of an extensive white matter network and showed that even late L2 learning leads to neural changes.

Concerning brain functions, these structural and functional changes have shown to affect both domain-specific functions such as language processing as well as domain general-executive functions.

Executive functions (EF), often referred to as cognitive control as a more general term, have been defined by Diamond (2013) as a “family of top-down mental processes needed when you have to concentrate and pay attention, when going on automatic or relying on instinct or intuition would be ill-advised, insufficient, or impossible” (Diamond 2013: 136). The most prominent categorization of EF according to Miyake et al. (2000) comprises inhibition as the ability to filter out irrelevant information as well as inhibiting dominant but unwanted responses (Bunge et al. 2002), switching (also called shifting) as the ability to shift between various mental task rules and refocus attention (Miyake et al. 2000) and working memory updating as the ability of holding information in the mind that is not visually present anymore but still relevant (Diamond 2013).

Since managing and therefore switching between multiple language seems to require both language control and executive control, a large body of behavioral literature has investigated potential beneficial effects of bilingualism on executive

functions such as task switching, inhibition, working memory updating, monitoring and attentional processing in bilingual children as well as adults and older adults compared to monolinguals. Bialystok et al. (2008) compared older and younger bilinguals with monolinguals in working memory performance, lexical retrieval, inhibition and attentional control and found that while bilinguals and monolinguals performed similar on working memory tasks, bilinguals outperformed monolinguals on executive functions tasks. Costa et al. (2008) and Costa et al. (2009) showed that bilinguals outperformed monolinguals in executive monitoring and inhibition, already highlighting the importance of task demands for an advantage in EF to surface. Poarch and Bialystok (2015) examined monolingual, partially bilingual, bilingual and trilingual children and found that bilinguals outperformed monolinguals in inhibition and monitoring, while there was no difference between partially bilinguals and monolinguals as well as trilinguals and bilinguals. Contrary to this, several behavioral studies showed no or only partially enhanced executive functions in bilinguals, highlighting research shortcomings such as inconsistent experimental design, task heterogeneity, participant size and task relevance and consistency (e.g. Paap and Greenberg 2013; Rosiers et al. 2019; van der Linden et al. 2018).

However, neuroscientific methods have offered a more fine-grained approach into cognitive processes and have heavily contributed to show the neural substrates of language switching and how they overlap with the network of general executive functions. In a PET study, Price et al. (1999) scanned bilinguals while they read or translated words in L1 and L2 or altered between L1 and L2. Results showed increased activation of the anterior cingulate cortex (ACC) and subcortical structures that are implicated in domain-general executive control. Using fMRI, Hernandez et al. (2000) investigated bilinguals in single-language and dual-language picture naming and found increased activation of the dorsolateral prefrontal cortex in the mixed language condition. In an EEG study Khateb et al. (2007) found electrophysiological evidence for overlapping language control and executive functions. Bilinguals completed monolingual task selection and bilingual language selection in picture naming paradigms which resulted in increased activation of left middle frontal precentral gyri, supramarginal and angular gyri. From this the authors concluded that language selection follows a network involved in domain-general executive control as well as language processing. Baene et al. (2015) used fMRI to investigate which brain regions overlap in activation during language switching and non-linguistic switching. Results showed that the prefrontal cortex (PFC), ACC and pre-supplementary motor area (pre-SMA) as well as the inferior and parietal lobe were active in both language switching and non-linguistic task switching.

Results from these studies show a fronto-parietal network that shows significant overlap with the network of different executive functions. Abutalebi and Green

(2007, 2008) mapped a first model that includes the left PFC, the ACC and pre-SMA, the basal ganglia and the bilateral inferior parietal lobule which as later expanded to include the right PFC, the thalamus, the cerebellum and the left putamen (Abutalebi et al. 2013; Abutalebi and Green 2016). This model partly overlapped with the findings of a meta-analysis by Luk et al. (2011) where they investigated the neural regions involved when bilinguals switch between languages. Regions that showed significant and consistent activation included the left inferior frontal gyrus, the left middle temporal gyrus, left middle frontal gyrus, right precentral gyrus, right superior temporal gyrus, midline pre-SMA and bilateral caudate nuclei. Due to this neural overlap, executive functions have been assumed to be implicated in several aspects of language processing such as coordinating lexical selection, preventing non-target interference, monitoring output and keeping the working memory in the feedback loop (Abutalebi and Green 2007: 249).

As research has shown, different factors such as age of acquisition of L2, language proficiency as well as language use and switching frequency influence the recruitment and change of the language and executive control networks (Calabria et al. 2018).

The executive function of switching (Miyake et al. 2000) also referred to as mental flexibility (Diamond 2013) holds a unique status, since it also includes aspects of other executive functions such as interference suppression and inhibiting the irrelevant response rule, updating the working memory with a new task and its rules as well as flexibly shifting between task sets (Cespón and Carreiras 2020; Jamadar et al. 2015). Switching or mental flexibility is often studied via task switching paradigms such as the color-shape switch tasks, the number letter tasks, or the Wisconsin Card sorting test (Grant and Esta 1948). Within these tasks, participants typically see single-task blocks that only contain repeat trials which means only performing one task rule as well as mixed-task blocks that contain both repeat and switch trials. In the mixed-task blocks, participants frequently must switch between task rules. In behavioral studies, switching and mixing costs are typically calculated based on reaction times. Switching costs are calculated by determining the RT difference between switch and repeat trials in mixed task blocks while mixing costs are calculated by comparing RT performance in single task block with performance on repeat trials in the mixed-task blocks (Braver et al. 2003).

Task switching was frequently compared directly with language switching to show how bilinguals can benefit from the neural overlap between language control and executive control and enhance their EF: Garbin et al. (2010) compared monolinguals to bilinguals in a task switching paradigm and found that bilinguals showed an advantage in switching depicted in reduced switching costs (see van den Noort et al. 2019 for a review). Declerck et al. (2017) also compared switching effects in linguistic and non-linguistic switching and found that when the stimuli, cues and

response modality were the same, language switching costs and task switching costs did not differ, however, when linguistic tasks were used, language switching and task switching costs differed significantly in error rates. They concluded that language switching, and task switching do share neural mechanisms to some extent, deducing that there is partial overlap between language control and domain-general executive functions.

Interestingly, as Prior and Gollan (2011) have shown by comparing monolinguals with two groups of bilinguals with different language switching frequencies, the amount of language switching has a significant effect on task switching advantage. Verreyt et al. (2016) replicated these results by comparing unbalanced bilinguals (BL), balanced non-switching bilinguals and balanced switching bilinguals in two non-linguistic switching tasks. While there was no difference in executive control performance in unbalanced BL and balanced non-switching BL, balanced switching bilinguals outperformed the other groups. As seen in Baene et al. (2015), switching between languages and switching between tasks involve overlapping neural areas. In a non-linguistic switching task, Garbin et al. (2010) found increased activation in the left inferior frontal cortex as well as the left striatum for bilinguals compared to monolinguals in switch trials, interpreting this enhanced involvement in determining the right response to a stimulus and enhanced inhibitory processes and highlighting the involvement of executive functions in language control. In another fMRI study, Weissberger et al. (2015) examined bilinguals in task switching and language switching and found a significant overlap in brain response in the bilateral thalamus, cingulate gyrus and right caudate for task switching and language switching with language switching showing a greater brain response. Interestingly, repeat trials in both language and task switching tasks showed greater activation in frontal, parietal, temporal, occipital, cingulate and subcortical areas, suggesting shared mechanisms of switching for both domains but also indicating more efficiency for keeping the inhibition of non-target stimuli for bilinguals. In an ERP-study, Declerck et al. (2021) compared language switching and task switching and found no significant differences in ERP patterns related to switching across language and task switching, supporting the assumption of EF function involvement in language switching.

Electrophysiological research on task switching and language switching frequently compared non-linguistic task switching in bilinguals and monolinguals by examining specific ERP components. Previous research in the context of task switching has differentiated between cue-locked ERPs that represent proactive control and prepare the system for a task switch and target-locked ERPs that represent reactive control that helps the system to overcome interference and switch attention to the relevant task (Jamadar et al. 2015). The ERP components N200 and P300 have been related to specific aspects of cognitive control during task switching (see Cespón and Carreiras 2020 for a review). The N200 is a fronto-central negative-

going component peaking around 200–350 ms after stimulus presentation that has been related to inhibitory processes (see Folstein and van Petten 2008 for a review) and attention switching. The N200 is assumed to reflect the effort of selecting the correct response in a mismatch or conflict situation, which is indicated by larger or more negative N200 amplitudes and delayed latencies for switch trials than for repeat trials (see Gajewski et al. 2018 for a review). In terms of processing demands, research on go/no-go trials that also require inhibition has associated more negative (larger) N200 amplitudes with increased activation of the inhibitory control system (Jodo and Kayama 1992).

The P300 component, subdivided into the P3b component relevant in this context, is positive going and peaks around 300 ms after stimulus onset. It is rather temporal-parietally distributed (see Polich 2007 for a review). In the context of task switching, the P300 has been related to updating the working memory to implement the relevant task rule as well as attention switching from the preceding to the latest stimulus-response mapping. Therefore increased working memory load and switching attention has been associated with smaller P300, which is why the P300 amplitude is usually smaller for switch than for repeat trials (see Cespón and Carreiras 2020; Gajewski et al. 2018 for a review). The easier the task, the larger the P300. These N200 and P300 modulations were observed in various task switching studies (see Jamadar et al. 2015 for a review) and specifically studied in the context of bilingualism such as by López Zunini et al. (2019). They compared monolingual and bilinguals and found larger N200 amplitudes for switch than for repeat trials and larger N200 for bilinguals than for monolinguals in both conditions showing that bilinguals exhibited enhanced conflict monitoring across all task conditions. Furthermore, they found that although repeat trials elicited larger P3b amplitudes than switch trials, bilingual older adults showed smaller P3b amplitudes in all conditions compared to monolingual older adults. Together with superior switching and mixing results of bilingual older adults, this suggested that bilinguals rely on different processing strategies than monolinguals as they age. Similarly, Timmer et al. (2017) found more negative N200 amplitudes for switch than for repeat trials in bilinguals, while monolinguals did not show any difference.

Several studies using either behavioral or neuroscientific methodologies or both in combination have highlighted the influence of processing demands as a requirement for enhanced EF (Costa et al. 2009; Dong and Liu 2016; Henrard and van Daele 2017; Pliatsikas et al. 2017; Prior and Gollan 2011). This means that high processing demands such as frequently occurring conflict must be present, as is the case in frequent language switchers. As a particularly taxing modality of switching between languages, simultaneous interpreting (SI) is a cognitively demanding task that rapidly gained interest in investigating executive functions both from a neurolinguistic as well as a translation studies perspective. During simultaneous interpreting, the interpreter

must simultaneously comprehend the input of new speech coming in, store it in memory, reformulate speech segments in their mind and monitor and articulate the output. Both languages need to be active at the same time while unintentional language switches must be avoided (Christoffels et al. 2006). SI has therefore been referred to as a modality of extreme language control (Hervais-Adelman and Babcock 2020). Accommodating these increased processing demands from managing multiple languages in the brain while performing overlapping processes of speech comprehension, translation and speech production is a skill that needs to be acquired. Behavioral results from translation studies that compare executive functions in translators and SI to bilinguals and or monolinguals have broadened the spectrum of the research field. Yudes et al. (2011) compared SI, monolinguals and bilinguals in tests of cognitive flexibility and inhibition and found that SI outperformed monolinguals and bilinguals on cognitive flexibility but not on inhibition. Dong and Liu (2016) compared translation students, interpreting students and bilinguals receiving L2 training before and after a 6 month training period in the EF of inhibition, switching and working memory updating and found that interpreting yielded significant advantages in switching and WM-updating while translation and L2 training did not produce significant advantages. Henrard and van Daele (2017) compared monolinguals with professional translators and SI in inhibition, cognitive flexibility and WM-updating and found that SI outperformed monolinguals in all EF and translators in all EF but cognitive flexibility. Babcock and Vallesi (2017) compared SI with multilinguals on conflict resolution, inhibition and task switching and found that although there was no significant difference between the groups in conflict resolution and switching costs, interpreters showed an advantage in mixing costs in task switching. van der Linden et al. (2018) compared professional simultaneous interpreters to monolinguals and L2 language teachers in inhibition, attention switching, short-term memory and WM-updating and found no difference in executive functions between the groups but a slight advantage in interpreters for short-term memory. In a systematic review, Nour et al. (2020) found that only the EF of switching and updating were improved in interpreters compared to different non-interpreting control groups and emphasized that especially switching abilities seem to be more sensitive to interpreting training and continue to develop during training. Referring back to Miyake and Friedman (2012) who stress that although different executive functions do overlap to some degree, they also have distinguishing features, Nour et al. (2020) further highlight how important it is to consider the diversity of executive functions and the multiple influencing factors such as the frequency and intensity of switching between languages.

Within a neurolinguistic approach, several studies have examined the neural correlates of executive functions in SI. In an ERP study, Elmer et al. (2010) tested SI as well as matched bilinguals on a auditory semantic judgment task and found increased an N400 ERP response for SI in the incongruent condition, indicating an

elevated sensitivity to semantic processing. Becker et al. (2016) used fMRI to compare professional SI and bilingual controls in the EF of switching and found increased grey matter volume in the left frontal pole in SI as well as increased connection of this region to the left inferior frontal gyrus and middle temporal gyrus compared to controls. Klein et al. (2018) examined resting state functional connectivity using EEG and compared SI to professional bilinguals and found hyperconnectivity between the ventral part of the prefrontal cortex and the dorsolateral prefrontal cortex (DLPFC) in the alpha frequency range for interpreters compared to controls, suggesting that SI leads to increased communication between executive control regions.

Since SI is an acquired skill that entails frequent language switching and controlling two languages under time pressure, several studies have investigated how SI trainees undergo plastic and functional changes in control-related networks during SI training. Hervais-Adelman et al. (2015) examined brain responses during SI and shadowing (speech repetition) in student SI at the beginning and after their training and compared them to matched controls. They found changes in the caudate nucleus that could be related to SI training. The caudate nucleus is a structure that has been identified as essential for executive functions and language control. van de Putte et al. (2018) compared simultaneous interpreting students with translation students before their respective training and after nine month of training, observing increased activation in the right angular gyrus and the left superior temporal gyrus in SI in during EF tasks of inhibition and switching. Furthermore, they found an increase of structural connectivity in the fronto-basal ganglia subnetwork related to domain-general executive functions as well as language control, and in a cerebellum and SMA subnetwork most prominently involved in language control. Dong and Zhong (2017) compared university students with more or less interpreting experience in a Flanker task within an ERP context. They found larger N1 and N2 amplitudes in congruent and incongruent stimuli, interpreting this as an advantage in attentional processing and monitoring. However, Cespón (2021) has emphasized that especially the N200 amplitude is often misinterpreted and larger N200 amplitudes have incorrectly been associated with an “advantage of bilinguals” while larger N200 amplitudes really rather relate to greater processing effort and increased resource allocation to inhibitory processes. It therefore seems as though acquiring and training the skill of simultaneous interpreting induces activity-dependent plasticity on both areas related to language control as well as EF. Interestingly, since some studies have also observed decreased grey matter volume in bilinguals compared to monolinguals (Elmer et al. 2011, 2014; Kaiser et al. 2015) and lower fractional anisotropy (FA) values for interpreters compared to controls (Elmer et al. 2011), interest has shifted towards a focus on the progression of language and cognitive control as a function of the experience of managing multiple languages. Activity-dependent plasticity (see Lövdén et al. 2013 for a review) seems to correspond to

concepts of skill learning, for example the concept by Chein and Schneider (2012), where the acquisition of a skill develops from a learning phase over a controlled execution phase to an automated execution phase. The latter does not tax cognitive control as much as in controlled execution or initial learning. If this is applied to the SI training experience, it is possible that learning to interpret simultaneously boosts the executive function of task switching initially, but that after a certain time of training and automation of executive function processes through SI training, performance in task switching reaches a plateau. This has been suggested in the bilingual domain by Paap (2018) as well as in a SI study by Dong and Liu (2016). Based on the model by Chein and Schneider (2012), Paap (2018) suggested the controlled dose hypothesis that assumes that L2 learners transition from a phase where they strongly rely on executive functions to a phase where they rely less on executive functions due to automation of cognitive processes supporting L2 learning and coordinating multiple languages. They assume that executive functions are boosted rapidly by the initially enhanced processing efforts and that after cognitive mechanisms dealing with the enhanced processing efforts have been automated, the boost abates (Paap 2018: 454). Dong and Liu (2016) have speculated that this also occurs for executive functions during SI training by assuming a development curve of executive control enhancement in SI that increases steadily at the beginning of training but might reach a peak and plateau once the SI competence has increased and therefore automation of certain processes requiring executive functions are automated.

Since being bilingual can be classified as one of the experiences that impact the structure and functionality of the brain, several models have been proposed to accommodate experience-dependent plasticity induced by various degrees of bilingualism. Grundy et al. (2017) have proposed the bilingual anterior-to-posterior and subcortical shift (BAPSS) hypothesis, where bilinguals eventually shift from recruiting frontal and executive regions to an increased recruitment of posterior and subcortical regions when solving tasks that involve executive functions. A similar approach is presented in the Dynamic Restructuring model (Pliatsikas 2020). It postulates an initial change in a network of parietal, temporal and frontal regions related to vocabulary, semantic and phonological learning, as well as EF due to an increased need to control lexical alternatives in Stage 1, which is reflected in increased grey matter volumes. Then, in Stage 2, superfluously formed connections are pruned due to increased immersion and bilingual experience, which results in a return to baseline volume. This does not mean that the bilingual brain has already automatized all control functions, but that the task in focus for immersed BL now is to continuously control between lexical alternatives, inhibit the non-target language and switch between languages as desired. This is reflected in a shift from focus on cortical structures in Stage 1 to subcortical structures of language control and EF. Stage 3 represents peak efficiency, where prior enhancements have renormalized for

optimal and automatized language and cognitive control. Pliatsikas (2020) emphasizes that the interpreting experience has such an intensity that normalization processes in Stage 2 and Stage 3 fuse together. Marin-Marín et al. (2022) found a non-linear relationship between a bilingualism score and grey matter volume in the inferior frontal gyrus (IFG), with reduced volume at the lower and higher end of the bilingualism spectrum. They interpreted grey matter reduction at the high end of the bilingualism spectrum in favor of an increased experience with switching between languages in a high-switching community, resulting in a decreased involvement of the IFG.

2 Research gap and hypotheses

As the literature reviewed above has shown, several studies have investigated how training and expertise development affect executive functions and brain structure as well as functionality. This was frequently done by using cross-sectional designs that only assess one point in time and do not represent the dynamic process of skill acquisition and activity-dependent plasticity. If longitudinal designs were chosen, they mostly occurred in the form of pre- and post-tests. Although the pre- and post-tests are a powerful methodology and are useful in detecting if changes occurred after the training has finished, they do not offer a detailed insight into the precise stage of training at which changes occur. Although structural changes and changes in functionality do not necessarily overlap, it is interesting to keep the findings on structural brain changes due to skill progression in mind. Paying attention to the findings that skill acquisition goes through different phases from controlled or supervised execution into automated execution, it is possible that skill automation is reflected by a return to almost baseline levels of different measures due to more efficient processing (Elmer et al. 2011; Marin-Marín et al. 2022). This is why the progress of skill development might not be well represented in pre- and post-tests.

Furthermore, behavioral measures to uncover how task switching is influenced by improved EF due to managing multiple languages have yielded homogenous results. Also, although some studies have used structural neurophysiological measures, electrophysiological evidence, especially for neural changes during interpreting training, remains scarce. However, electroencephalography (EEG) and the ERP technique are powerful tools to gain insight into the temporal dynamics of changes in the brain, offering an extremely high temporal resolution of brain reactions to a stimulus.

In the light of the above, I conducted a longitudinal EEG study that goes beyond pre- and post-test but also records the “in-between.” Within the scope of a doctoral dissertation, the enhancement of executive functions through interpreting training,

more specifically the EF of inhibition, switching and working memory updating (Miyake et al. 2000), is investigated. However, within the scope of this paper, the focus lies on the executive function of switching which is investigated by administering a task-switching paradigm. The executive function of switching entails aspects of interference suppression as well as aspects of attention allocation and working memory updating in the form of stimulus categorization (see Cespón and Carreiras 2020; Jamadar et al. 2015).

Research on professional as well as trainee SI suggest a change in the EF of switching related to the increased task demands of SI as an “extreme” form of language control (see Hervais-Adelman and Babcock 2020). Therefore, as a part of my doctoral dissertation, I am interested in providing electrophysiological longitudinal evidence for a change in the EF of switching over the course of an interpreting training program. I expect that interpreting training, since it taxes both language control and executive functions, enhances task switching up until a certain point of automation, after which a ceiling effect is reached, which is reflected in seemingly decreasing ERP correlates for task switching.

For the ERP components, I expect switch trials will show significantly more negative (bigger) N200 amplitudes than repeat trials based on relevant ERP-literature in task switching (see e.g. Declerck et al. 2021; López Zunini et al. 2019; Timmer et al. 2017) since they require more interference suppression from the previous task rule than repeat trials and more resource allocation to interference suppression has been associated with more negative N200 amplitudes (see Cespón and Carreiras 2020; Kopp et al. 1996). Furthermore, I expect the N200 to become less negative over the course of testing for both conditions since interpreting training has been associated with improved interference suppression (Dong and Zhong 2017) and the N200 amplitude has shown to be less negative when less resources for inhibitory effort need to be allocated (see Cespón and Carreiras 2020; Kopp et al. 1996). For the P300 component, I expect that switch trials will show significantly less positive (smaller) P300 amplitudes than repeat trials based on relevant ERP task-switching literature such as by López Zunini et al. (2019), Periañez and Barceló (2009) or Chen et al. (2022) (and see Cespón and Carreiras (2020) for a review), since smaller P300 amplitudes are associated with more allocation of attentional resources (Polich 2007) and switch trials require more attentional resources since they involve the switching of attention between the stimulus representation and the corresponding response mapping during task-set reconfiguration (Monsell 2003). Correspondingly, I expect the P300 amplitude to increase over the course of testing since SI training has shown to improve task switching abilities (see Babcock et al. 2017; Dong and Liu 2016) rooted in attention switching (Monsell 2003), and allocating less cognitive resources to attention switching and stimulus categorization has been associated with increased P300 amplitudes (Cespón and Carreiras 2020). I also hypothesize that the progression

of the voltage in the respective time-windows will reach a plateau at a certain point in training after attention reallocation and interference suppression that are trained during SI training are automated to a certain degree as proposed in relevant bilingualism literature (see Chein and Schneider 2012; Paap 2018) as well as SI studies (Dong and Liu 2016).

I hypothesize that participants in the experimental group and in the control group do not differ in the task switching abilities at baseline based on literature by Babcock et al. (2017), Dong and Liu (2016) and Rosiers et al. (2019) who did not find a behavioral difference between SI and other control populations at baseline testing.

3 Methods

In total, 24 students (2 male, 20 female) participated in the EEG-study. 10 students from the MA Conference Interpreting training program with the language combination of German-English were tested on the executive functions of task switching before, during and after their 2-year SI training (4 Semesters). A Control group consisting of 14 MA Translation students was also tested prior to the beginning of their MA Translation. The Control group was only tested once. This is because the longitudinal study started in the middle of the Covid-19 pandemic in November of 2021 when lab use was restricted. Therefore, the decision was made to follow through with the interpreting group and test a Control group at a later point in time to compare baseline performance.

All participants took part in the study voluntarily and gave their informed consent and were informed that they could withdraw their participation from the experiment without any consequences at any given time.

3.1 Participants

10 SI students and 14 translation students participated in the study. Data from two Control group students was excluded from data analysis since they had some interpreting experience after all. Before taking part in the first EEG session, participants filled out the TICQ (Translation and Interpreting Competence Questionnaire) (Schaeffer et al. 2020), a validated questionnaire consisting of different sections tapping into general bilingual proficiency and interpreting- and translation-related proficiency. The TICQ offers the possibility to extract the raw answers from the questionnaire as well as to obtain a summed competence score per participant consisting of weighted values of the given answers. During its validation process, the robustness of the different sections of the TICQ to discriminate between lay

bilinguals, students and professionals was tested by using multigroup discriminant function analysis (MDA) which revealed two discriminant functions F1 and F2. These functions were able to successfully discriminate between groups and to predict the respective group membership. The summed competence scores per participant represent the F1 and F2 values (Schaeffer et al. 2020: 97–100). Increased F1 values have been associated with increased self-rated competence values (Jacob et al. 2024: 552) and the F1 reflects individual skill associated with the translation process (Jacob et al. 2024: 548–549).

Participants were matched in age with a mean age of participants in the Experimental group (SI group) of 24 years ($SD = 2.5495$) and 23.92 years ($SD = 3.629$) for the Control group. As indicated by a paired t -test, there was no significant age difference between groups ($p = 0.9538$). All participants in the SI group grew up in Germany and indicated German as their L1 and English as their L2 except one participant who indicated English as their L3. 9 of 12 participants in the Control group had English as their L2, 3 spoke English as their L3. Except for one participant in the Control group, none of the participants in the SI group or in the Control group were fluent in English before the age of seven. Participants were matched on age of English acquisition, with mean age of English acquisition in the SI group of 8.33 years ($SD = 2.06$) and mean AoA of 8.08 years ($SD = 2.7784$) in the Control group. A paired t -test showed that there was no significant difference between groups in terms of English AoA ($p = 0.823$). Furthermore, all participants learned the L2 in the context of formal education and all participants indicated to speak English between 1 and 15 h per week. Before starting their respective programs in translation and interpreting, participants in the SI group rated their own English proficiency on a scale from 1–100 on average with 76.11 ($SD = 12.244$) and the Control group rated their English proficiency on average with 75.833 ($SD = 10.624$). There was no significant difference in self-rated English proficiency between groups ($p = 0.958$).

On the same scale, participants in the SI group rated their proficiency in interpreting from their L1 into their L2 on average as 44.44 ($SD = 19.5966$) and their proficiency in interpreting from their L2 into their L1 on average as 61.666 ($SD = 18.2000$). Participants in the Control group rated their proficiency to interpret from their L1 into their L2 with a mean of 37.25 ($SD: 32.30$) and their proficiency to interpret from their L1 into their L1 with a mean of 44.75 ($SD = 34.83$). Although the means of the SI group was larger for both directions at baseline, a t -test did not indicate a significant difference between groups in the L1–L2 direction ($p = 0.535$) or in the L2–L1 direction ($p = 0.168$). All participants either graduated with a BA of translation, BA in linguistics or education. None of the participants had any immediate prior substantial interpreting experience before participating in the study. Two participants in the Experimental group were excluded after the third session since they quit their SI training program. ERP-data from one participant had to be excluded

in Session two due to data loss. In Session 4, some participants spent an Erasmus semester abroad, which is why one participant could not be tested. However, after this participant returned and continued their SI training, they were reincluded for the last session.

3.2 Stimuli and procedure

Before the EEG experiment started, participants filled out a consent form as well as the TICQ. For the EEG study, participants were seated in an EEG chamber isolated as a Faraday cage in front of a Desktop and keyboard. Participants were instructed to avoid movements during the EEG recording and were informed to blink as little as possible and to keep their jaw relaxed. Participants completed a Flanker task to test inhibition, a color-shape switch task to test task switching as well as a n-back task to test working memory updating in a randomized order. The EEG study consisted of 5 sessions. The first session took place before participants started their training and functioned as baseline. After baseline testing, participants were tested at the end of every semester for four semesters.

The color shape-switch task (Meiran 1996) consisted of circles and triangles in blue and red as well as a color and a shape cue that indicated the task for the following stimulus. Stimuli presentation was computed with PsychoPy (Peirce 2007). Metrics and stimuli number were based on literature and adapted from López Zunini et al. (2019) and Prior and Gollan (2011). In repeat trials, participants saw only one type of cue and had to complete only one task (e.g. react to shape of stimulus). In switch trials, the task switched as indicated by the cue (e.g. from react to shape to react to color). In an initial training block consisting of either one block per task in the repeat condition or one training block for the switch condition, participants familiarized with the task. In the actual experimental task, participants saw 2 single-task blocks containing 56 trials each and 4 blocks mixed-task blocks, each containing 56 trials each. Stimuli and cues were pseudo-randomized with Python. Before each cue, participants saw a fixation cross for 300–350 ms followed by a black screen for 150 ms. The respective cue was then presented for 250 ms. Time of stimulus was presented for 4,000 ms or disappeared when a participant pressed a key. The stimulus was followed by an intertrial interval of 850 ms. If a shape cue was shown, participants had to use their left hand and press the key “S” for circles and “D” for triangles. If a color cue was shown, participants had to use their left hand and press the keys “L” for blue and “K” for red.

In total, one experiment session lasted 1 h (time to place the electrodes excluded), the color-shape switch task itself covered about 20 min.

3.3 EEG recording and data processing

A continuous Electroencephalogram EEG was recorded by an elastic electrode cap with 64 Ag/Ag + electrodes placed on the scalp in accordance with the International 10/20 system, using NeuroScan Synamps2 (Compumedics, El Paso, TX, USA) with a sampling rate of 1,000 Hz. Linked mastoids were averaged for referencing and impedances were kept below 5 k Ω . EEGLAB (Delorme and Makeig 2004) was used for off-line processing, where data was down sampled to 250 Hz in order to process data faster. A high-pass filter at 0.5 Hz and a low-pass filter at 40 Hz was applied to filter the signal. After manual channel rejection of bad channels, an independent component analysis (ICA) was performed. Subsequently, artifacts were diminished by rejecting noisy ICA components manually. Finally, interpolation replaced the rejected channels and bad epochs were manually rejected. Epochs of 100 ms before stimulus onset and 700 ms after stimulus onset.

Incorrect trials and trials without an answer were rejected both from the ERP as well as from the RT analysis.

3.4 Data analysis

For the main analysis of behavioral data as well as the ERP analysis, linear mixed models were constructed in the RStudio environment (Posit team 2024) using the lme4 package (Bates et al. 2015). The linear mixed models for the behavioral data included reaction time as a response variable and Session as well as Condition (switch or repeat) as predictors. Additionally, SubjectID was included as random effect. For the ERP analysis, several electrodes were summarized as AOI (areas of interest), based on reviewed literature (see e.g. Declerck et al. 2021; López Zunini et al. 2019; Timmer et al. 2017).

Linear mixed models with Voltage in the specialized ERP time windows as response variable and Session as well as Congruency as predictors were fitted. SubjectID was included as random effect. Since multicollinearity can adversely affect the results of the mixed regression models, a variance inflation factor (VIF) was calculated to measure whether there is multicollinearity between the different independent variables, meaning that two more more variables are too highly correlated. Following Zuur et al. (2010), VIF-values that are equal to 1 indicate that there is no collinearity while VIF-values equal or larger than 3 were taken as evidence for high multicollinearity (Zuur et al. 2010: 9).

Furthermore, a Cohen's f^2 was calculated to examine the effect sizes of the independent variables. In contrast to Cohen's d that is not recommended for linear mixed models since it cannot accommodate the different sources of variance in

linear mixed models, the Cohen' f^2 has been applied frequently in linear mixed regression modelling (Selya et al. 2012: 1). According to Cohen's (1977) explanation, f^2 -values of 0.02 represent small effect sizes, values of 0.15 represent medium effect sizes and f^2 -values of 0.35 represent large effect sizes (Cohen 1977: 413–414).

4 Results

In the following subsections, behavioral as well as electrophysiological results are presented. Results are discussed and put into context in 5. Discussion.

4.1 TICQ

Matching English proficiency of the Experimental group in the pre- and post-questionnaire showed that although mean English proficiency increased from 76.11 (SD = 12.44) to 78.75 (SD = 10.26) there is no significant increase in self-rated English proficiency from baseline to the post-test ($p = 0.6806$). Self-rated proficiency of interpreting from L1 into their L2 increased from 44.44 (SD = 19.6) to 58.38 (SD = 31.1), however, a t -test did not indicate a significant difference between pre- and post-questionnaire ($p = 0.4128$). The other interpreting direction from L2 into L1 increased from a mean of 61.67 (SD = 18.20) to a mean of 75.25 (SD = 19.17). A t -test did not show a significant difference between pre- and post-questionnaire ($p = 0.2805$). When comparing the summed competence scores calculated per participant from pre- and post-training questionnaire, results showed a mean value of 4.297 (SD = 0.919) for the pre-test and a mean value of 5.116 (SD = 0.881) for the post-test. A t -test indicated a significant difference between pre- and post-training competences values ($t = 1.8945786$, $df = 7$, $p = 0.0415$). A Cohen's d for effect size revealed a strong positive effect of 0.9098, pointing towards a significant increase in SI competence.

4.2 Behavioral results

After fitting a linear mixed model with reaction times (RT) as dependent variable and group (Experimental vs. Control) and Condition (switch vs. repeat) as independent variables as well as SubjectID as random effect, overall reaction times (RT) at baseline (Session 1) did not differ significantly between experimental and Control group ($p = 0.984$, $SE = 1.078e-01$), indicating no prior cognitive advantage of SI over translators (see Figure 1). Switch trials showed significantly longer reaction times than repeat trials ($p < 0.01$, $SE = 1.044e-02$), this effect was small as indicated by

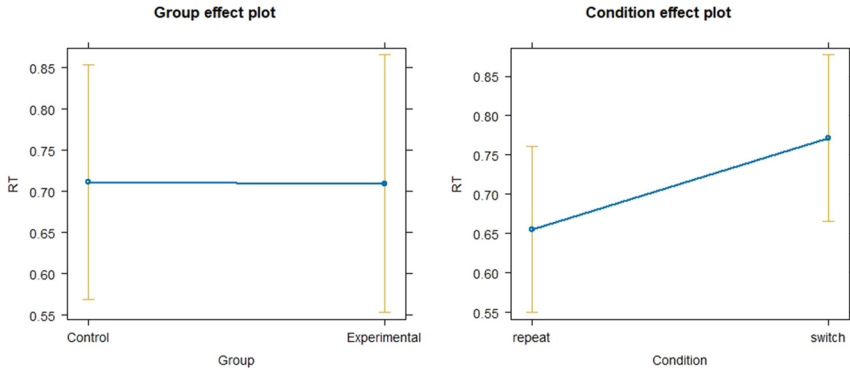


Figure 1: Linear mixed model of reaction times in experimental and control group at baseline.

$f^2 = 0.02$. Residuals showed that the data deviated from normality. Therefore, a log transformed linear model was fitted. The log transformed model did not reveal substantially different results compared to the non-log transformed model (effect of Group: $p = 0.887$; $SE = 0.135$; effect of Condition: $p < 0.01$; $SE = 0.010$). For this model, the VIFs for Group and Condition were equal to 1, indicating a low risk of collinearity.

Adding Condition as an interaction term did not reveal a significant interaction ($p = 0.075$, $SE = 0.02106$).

For accuracy, a logistic regression model was fitted with Accuracy as response variable and Group and Condition as predictor variables. Results revealed a small ($f^2 = 0.05$) but significant effect of Group ($p < 0.01$, $SE = 0.1033$) indicating that the Experimental group made more errors at baseline than the control group. The effect of Condition on Accuracy was small ($f^2 = 0.03$) but significant ($p < 0.01$, $SE = 0.1024$) with more errors made in the switch condition at baseline. VIF for Group and condition was equal to 1, indicating a low risk of multicollinearity. By adding Condition as an interaction term to check for an interaction between Group and Condition, a significant result was revealed ($p = 0.0394$, $SE = 0.2166$), the effect size of the interaction was small ($f^2 = 0.02$). The effect seems to be driven by the switch condition, indicating lower accuracy in the switch condition for the experimental group (see Figure 2).

Concerning the longitudinal behavioural data of the SI group, mean RT from both switch as well as repeat trials across sessions steadily decreased, with switch trial mean RT always being slightly longer as depicted in Figure 3.

After fitting a linear model with RT as dependent variable and Session (1–5) and Condition (switch vs. repeat) as independent variables as well as SubjectID as random effect, results showed significant effects for all sessions (Session 2: $p < 0.01$, $SE = 9.771e-03$, Session 3: $p < 0.01$, $SE = 9.461e-03$, Session 4: $p < 0.01$, $SE = 1.054e-02$,

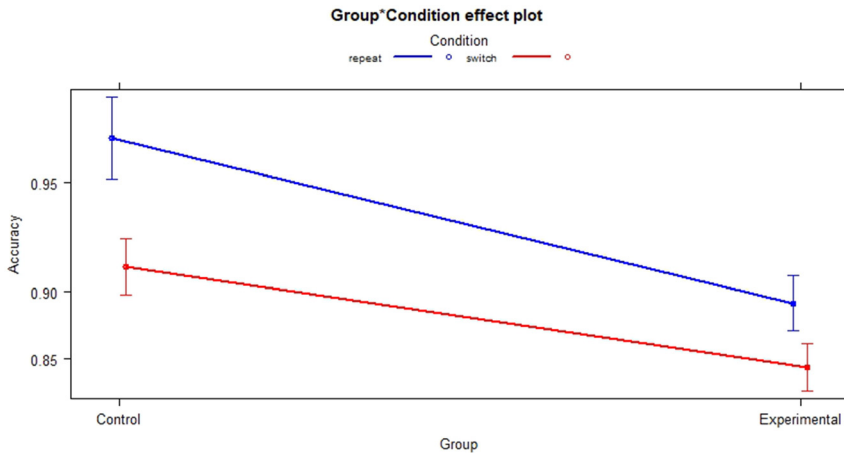


Figure 2: Interaction group and condition at baseline.

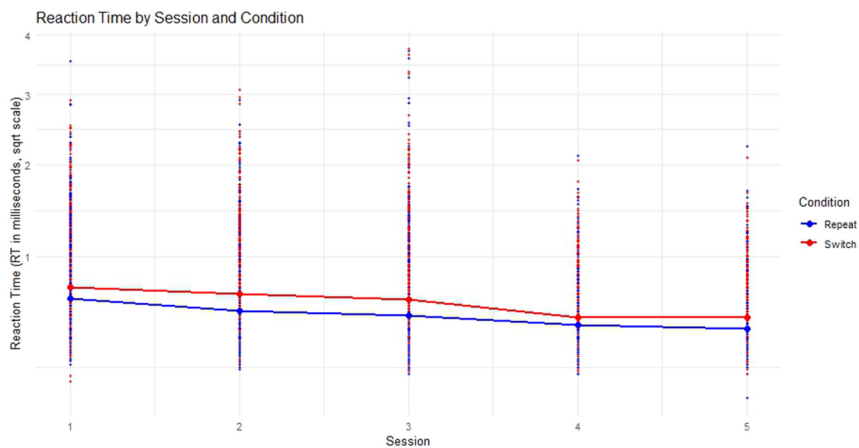


Figure 3: Mean reaction times across session for switch and repeat trials.

Session 5: $p < 0.01$, $SE = 1.018e-02$). The general effect size of Session was small as indicated by $f^2 = 0.07$. Furthermore, Condition showed a small ($f^2 = 0.02$) but highly significant main effect on reaction time ($p < 0.01$, $SE = 6.270e-03$) with repeat conditions showing significantly faster reaction times than switch conditions as seen in Figure 4. Residuals were not normally distributed. To remedy this, the fitted linear model was log transformed, however, log transformation did not lead to meaningful changes in the model outcome (all p -values < 0.01). The VIF-values were made

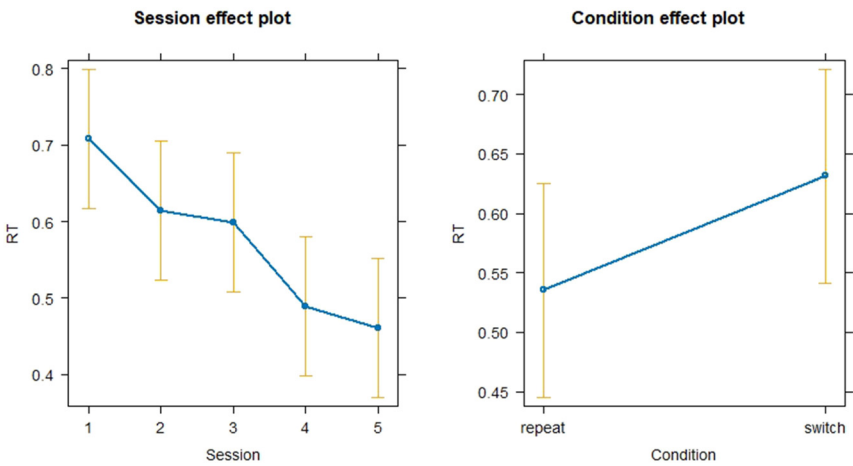


Figure 4: Linear mixed model for the effect of session and condition on reaction time.

comparable across all levels and were subsequently equal to 1 for both Session and Condition which indicated that there is no multicollinearity between independent variables.

As for interactions depicted in Figure 5, the linear mixed model showed a significant and positive interaction between Session 2 and the switch condition (0.0431, $p = 0.025$) which indicates that the RT for the switch condition in Session 2 decreases

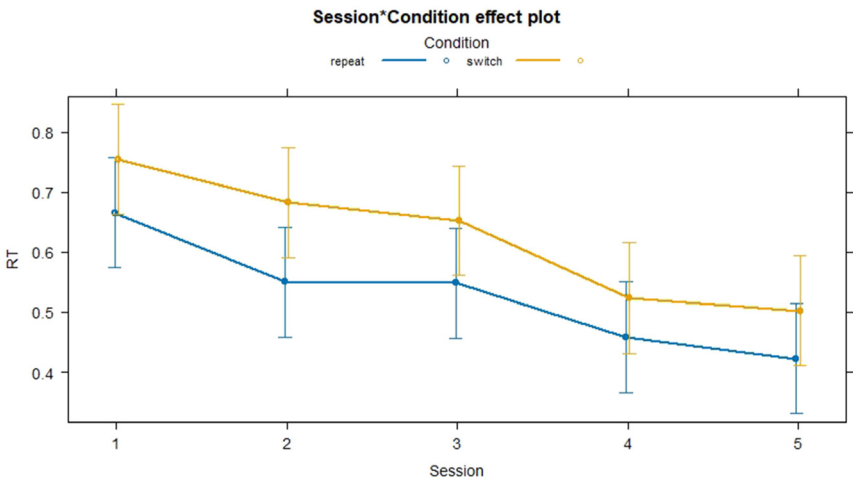


Figure 5: Interactions reaction times switch and repeat conditions.

significantly less than the repeat condition. The other sessions did not reveal any significant interactions with Condition, showing that the difference in RT between conditions remains consistent from Sessions 3 to 5 compared to Session 1. The effect size of the Session:Condition interaction was extremely small ($f^2 < 0.01$).

Behavioral switching costs were calculated based on reaction time differences of switch and repeat trials in the mixed task blocks. As can be seen in Figure 6, switching costs initially increased in Session 2 and then steadily decreased until a negative peak at Session 4, after which they increased again. A linear mixed model was fitted with switch costs as dependent variable, Session as predictor and Participant as random effect. The model did not reveal a significant effect of any Session switch costs (all $p > 0.05$).

The overall accuracy steadily increased except for a plummet at Session 3 after which it increases again for Session 4, followed by a slight decrease as depicted in Figure 7.

A logistic regression model was fitted with Accuracy as the outcome variable and Session and Condition as predictor variables. The model revealed highly significant main effects of all Sessions on Accuracy (S2: $p < 0.001$, SE = 0.12685; S3: $p < 0.001$, SE = 0.10374; S4: $p < 0.001$, SE = 0.15147; S5: $p < 0.001$, SE = 0.12766). The main effect size of Session was small ($f^2 = 0.08$). Furthermore, the model revealed a small ($f^2 = 0.03$) but significant main effect of Condition on Accuracy ($p < 0.001$; SE = 0.08301) with significantly lower accuracy for the switch condition. The VIF-values for Session and Condition were equal to 1 and did not indicate any sign of multicollinearity.

An interaction term was added to check whether Condition is modulated by Session. A significant interaction of Session 5 with condition was revealed ($p < 0.01$, SE = 0.29822) with the switch condition showing a decreased accuracy compared to

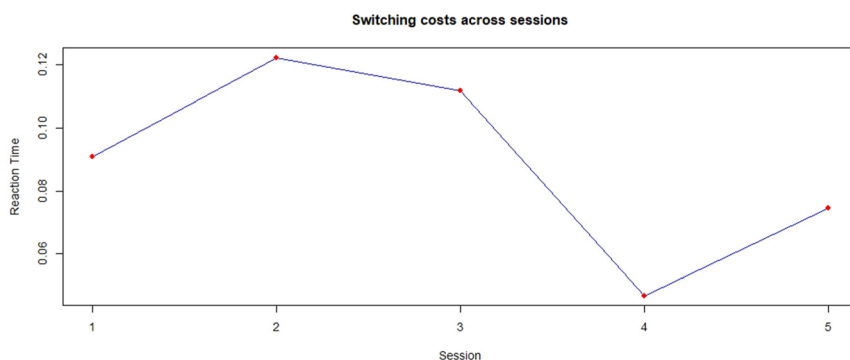


Figure 6: Behavioral switching costs across sessions.

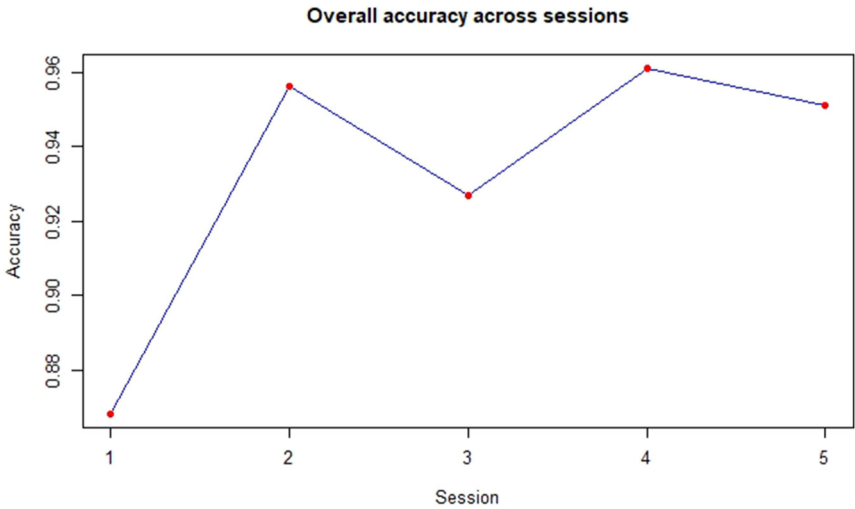


Figure 7: Overall accuracy for the interpreting group.

the repeat condition (see Figure 8). Overall, the effect size of the Session:Condition interaction was small as indicated by $f^2 = 0.03$.

To account for a possible speed-accuracy tradeoff, a balanced integration score (BIS) was calculated as presented in Lisefeld and Janczyk (2019). The BIS gives equal measures to reaction times and proportion of correct responses (PC) and is calculated by bringing reaction times and accuracy to the same scale through standardization

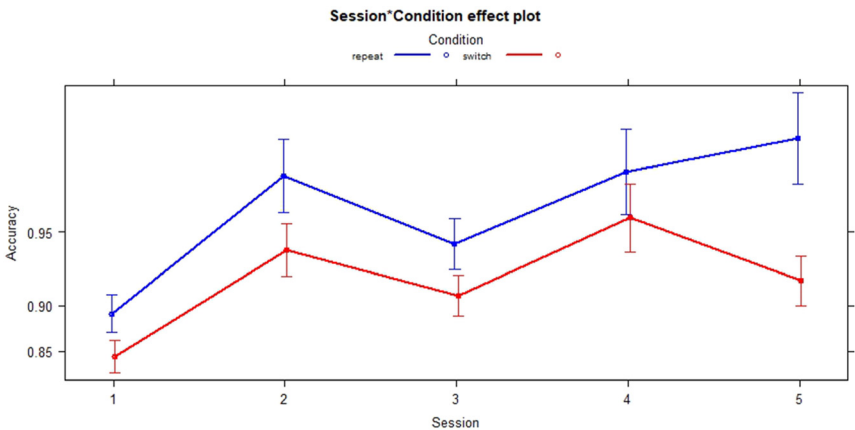


Figure 8: Interactions of session and condition across sessions.

and subsequently subtracting the standardized reaction time from the standardized PC.

$$BIS = Z_{PC} - Z_{RT}$$

This is done by calculating all measures across all subjects and conditions. The BIS represents therefore the difference between the standardized mean correct reaction times and the standardized PC (Lisefeld and Janczyk 2019: 42). Since the current data stems from a longitudinal study design, each session was treated as a separate experiment when calculating the PC. However, PC and RT were standardized across all sessions and conditions. Figure 9 depicts the progression of the mean BIS across all Sessions.

The BIS is calculated by subtracting the standardized RT from the standardized PC, which is why smaller Z_{RT} indicate faster responses than average and larger Z_{PC} indicate higher accuracy than average. A positive BIS value therefore is created when smaller Z_{RT} values are subtracted from larger Z_{PC} values. The increase in BIS values depicted in Figure 9 therefore indicates that participants increased their speed without trading in much accuracy, therefore becoming more efficient with SI training. To see whether Z_{RT} really decrease over Sessions and Z_{PC} really increase of Sessions, leading to increasing BIS values, the mean Z_{RT} and mean Z_{PC} were plotted as depicted in Figure 10.

To investigate whether Session and Condition have an effect on the BIC, a linear mixed model with BIC as dependent variable, Condition and Session as predictor variables and Participant as random effect was fitted. Residuals indicated that data

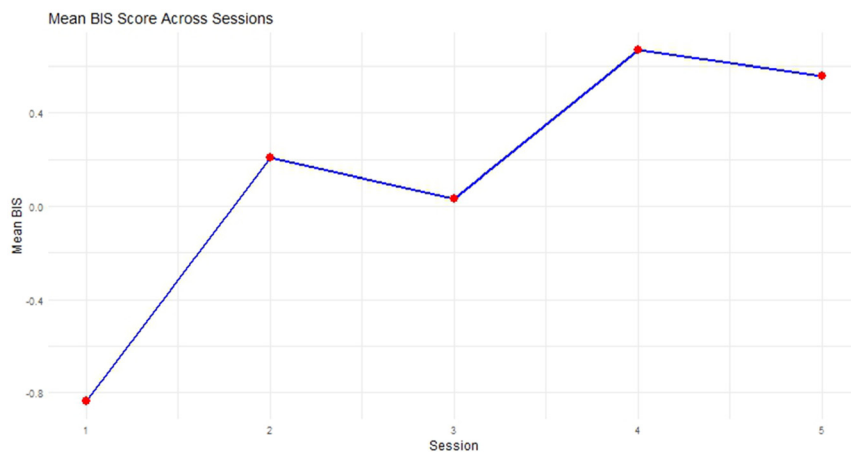


Figure 9: Mean balanced integration score per session.

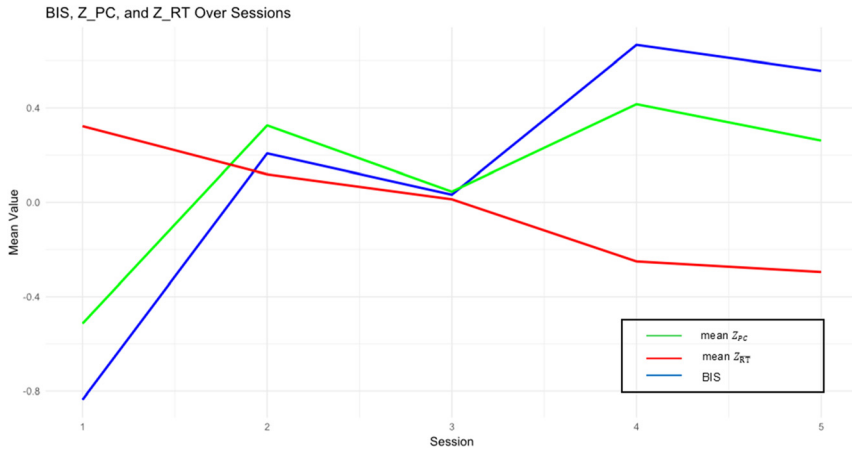


Figure 10: Progression of mean BIS, standardized mean RT and standardized PC across sessions.

was not normally distributed which is why a log-transformed linear model was fitted. Results from this model revealed a small to medium-sized ($f^2 = 0.07$) and highly significant main effect of all Sessions (S2: $p < 0.001$, SE = 0.036, S3: $p < 0.001$, SE = 0.035, S4: $p < 0.001$, SE = 0.036, S5: $p < 0.001$, SE = 0.036). Furthermore, the effect of Condition on BIC was small ($f^2 = 0.01$) but highly significant ($p < 0.001$, SE = 0.019) with the fixed effects coefficient of the switch condition decreasing by -0.15867 units, indicating that the switch condition was characterized by a lower BIC than the repeat condition. The VIF-value for Session was 1.4 and the VIF-value for Condition was 2.17, both values still in the acceptable range for collinearity indicated by Zuur et al. (2010). Adding an interaction term showed a trend towards significance for the interaction of Condition with Session 2 ($p = 0.066$, SE = 0.058), a significant interaction between Condition and Session 3, 4 and 5 (S3: $p = 0.02$, SE = 0.07; S4: $p = 0.03$, SE = 0.07; S5: $p = 0.003$, SE = 0.07). The decreasing coefficients of the switch condition indicate that these interactions are driven by the switch condition. The VIF-value for the Session condition interaction lays at 1.6, still in the acceptable range for collinearity.

4.3 Results from event-related potentials

For the SI group which was tested at Baseline and at the end of every Semester for four semesters, event-related potentials were analysed in two time windows that were based on literature commenting on ERP components in task switching paradigms. To investigate the P300 component, the time window of 280–400 ms after

stimulus onset was chosen, for the N200 component, the time window of 200–320 ms after stimulus onset was investigated.

4.3.1 Time window 280–400 ms

To investigate the P300 component, an area of interest based on ERP literature of the electrodes cz, cp1, cpz and p1 within the time window of 280–400 ms after stimulus onset was set.

Voltages within the P300 time-window for the SI and the Control group were compared at baseline level by fitting a linear mixed model with Voltage as dependent variable, Group and Condition as predictor variables and Participant as a random effect. Data showed a normal distribution. A small- to medium-sized ($f^2 = 0.1$) and significant effect of Group was found ($p = 0.0393$, $SE = 0.3389$) with the Control group showing significantly lower voltages than the Experimental group. Since lower P300 voltages are associated with more cognitive control effort (Cespón and Carreiras 2020; Gajewski et al. 2018), the SI seem to show an initial advantage in EF. However, looking at the behavioral data, this was accompanied by significantly larger accuracy for the Control group (0.928, $SD = 0.239$) than for the Experimental group (0.868, $SD = 0.338$), showing that the interpreters made significantly more errors at baseline than the translators ($p < 0.01$, $SE = 0.10289$). The effect of Condition was small ($f^2 = 0.04$) and not significant ($p = 0.1969$, $SE = 0.3375$). The VIF values for Group and Condition were equal to 1, excluding the possibility of multicollinearity between variables.

Adding Condition as an interaction term did not reveal a significant interaction between Group and Condition ($p = 0.571$, $SE = 0.683$).

For the separate sessions of the SI group, the following graphs revealed amplitudes that can be interpreted in terms of the P300 component (see Figure 11).

The progression of mean voltage for the SI Experimental group across sessions varied as shown in Figure 12. Repeat trials always showed higher mean voltage than the switch trials, both conditions peaking at Semester 4 after which voltage declines again almost to baseline level in the switch condition.

For the progression of the switch costs in the P300 time-window, a linear mixed model was fitted with switch costs per participant as dependent variable, Session as predictor and Participant as random effect. The linear model did not reveal any significant results of Session on switch costs (all $p < 0.05$).

For the progress of the SI group, another linear mixed model was fitted with Voltage as response variable, Session and Condition as predictors and SubjectID as random effect. Data was normally distributed. Results can be viewed in Figure 13. For Session 2, although voltage increases by 0.6029 units compared to Baseline, the effect does not reach significance ($p = 0.08058$, $SE = 0.3403$). For Session 3, voltages increases

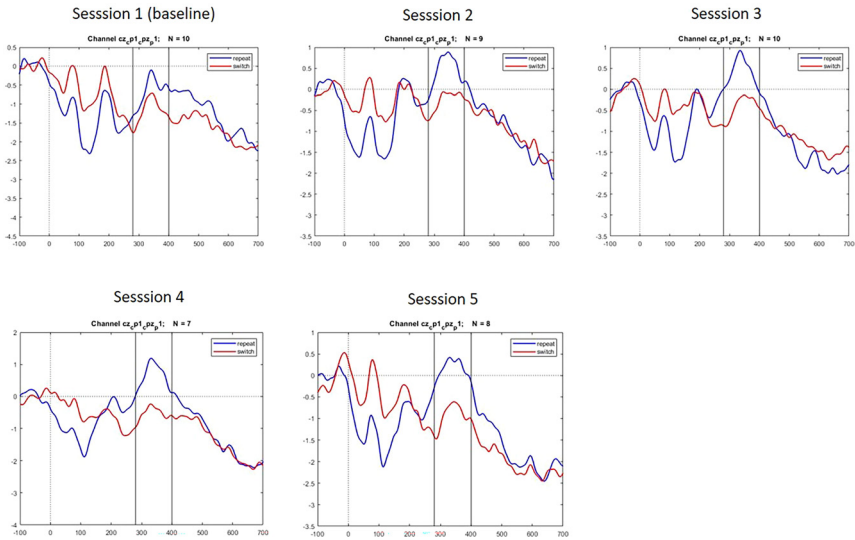


Figure 11: Event-related potentials in the P300 time-window across sessions.

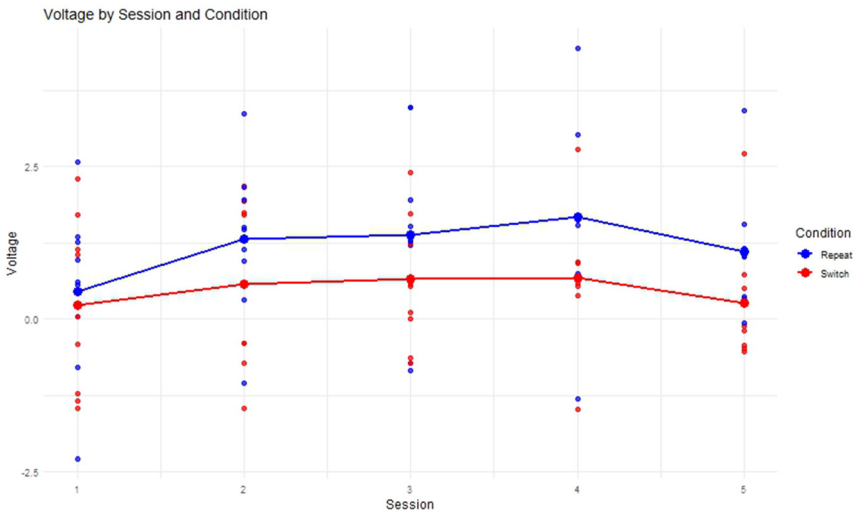


Figure 12: Progression of mean voltage in the P300 time-window per session.

by 0.6862 units compared to Baseline and this effect is statistically significant ($p = 0.04076$, $SE = 0.3295$). For Session 4, voltage increases by 0.7825 units compared to Baseline, showing a statistically significant effect ($p = 0.03796$, $SE = 0.3704$). For

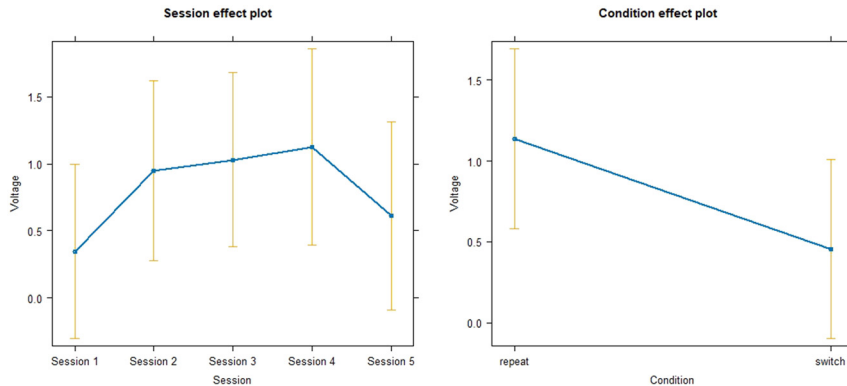


Figure 13: Linear mixed model for the effect of session and condition on voltage in the P300 time-window.

Session 5, the last session, Voltage increases by 0.2688 units compared to Baseline, the effect is not significant ($p = 0.45043$, $SE = 0.3543$). Overall, the variable of Session shows a small effect as indicated by the Cohen's f^2 ($f^2 = 0.05$). The variable Condition shows a small ($f^2 = 0.07$) but significant effect ($p = 0.00308$, $SE = 0.2221$), voltage decreases by 0.6799 units in the switch condition compared to repeat condition (used as baseline). The VIF for both Session and Condition were made comparable across levels and were consequently equal to 1, indicating that the independent variables do not show multicollinearity.

After including an interaction term of Condition into the linear mixed model see how the effect of Session on Voltage changes depending on Condition, the following results were generated. There was no significant effect of Condition on Voltage ($p = 0.6303$, $SE = 0.4746$) and no significant interaction terms of Session (all $p > 0.05$), suggesting that the effect of Session on Voltage is not significantly altered by Condition.

To see whether BIS is a factor that influences the voltage in the P300 time window, a linear mixed model with Voltage as dependent variables, Session, Condition and BIS as predictors and Participant as random effect was fitted. Data was normally distributed. Results revealed significant effects on Voltage for Session 2,3 and 4 (S2: $p = 0.005$, $SE = 0.298$; S3: $p = 0.007$, $SE = 0.269$; S4: $p = 0.003$, $SE = 0.309$) and marginally significant effect of Session 5 on voltage ($p = 0.058$, $SE = 0.290$). Furthermore, results showed a highly significant effect of Condition ($p < 0.001$, $SE = 0.166$). However, the effect of BIS on voltage was not significant ($p = 0.61$, $SE = 0.12$). The VIF values for Session, Condition and BIS all lay in the acceptable range for collinearity (1.05, 1.02, 1.25). To see whether the effect of Condition on P300 differs by BIS, the change across Sessions depends on BIS or whether the whole Session:Condition effect

changes with BIS, interaction terms were added. However, none of the three-way interactions were significant, so the effect of BIS on P300 voltage does not change in a different way across the sessions or between the switch and repeat conditions. There was merely a significant Condition: BIS interaction ($p = 0.02$, $SE = 0.18$) that was driven by the switch condition as its coefficient decreases by -0.4125 units, indicating that as BIS increases, the difference between switch and repeat trials becomes greater since switch trials become more negative. However, Session does not seem to impact this relationship.

4.3.2 Time window 200–320 ms

To investigate a possible N200 component, an area of interest based on ERP literature of the electrodes *fz*, *fcz* and *cz* within the time window of 200–320 ms after stimulus onset was defined.

For the time window 200–320 ms after stimulus presentation, a linear mixed model with Voltage as response variable, Group and Condition as predictors as well as SubjectID as random effect was calculated to account for possible differences between groups. Data was normally distributed. Results revealed a small ($f^2 = 0.09$) but significant effect of Group ($p = 0.00779$, $SE = 0.3957$) with a decrease of voltage by 1.1245 units for the Control group compared to the Experimental group, showing that the voltage of the Control group was significantly lower at Baseline compared to the Experimental group. Since according to N200 literature, a more negative N200 is associated with more cognitive effort or more executive control processes (Jodo and Kayama 1992), this could indicate towards an initial advantage of the interpreting trainees. Again, together with the increased error rate of the interpreting group over the Control group, the indication of an initial advantage is not clear cut. The main effect of Condition on voltage was extremely small ($f^2 = 0.002$) and not significant ($p = 0.559$, $SE = 0.382$). The VIF for Group and Condition was equal to 1, revealing no sign of multicollinearity.

Adding an interaction term did not show significant results ($p = 0.353$, $SE = 0.768$).

For the separate sessions of the SI group, the following event-related potential graphs revealed amplitudes that can be interpreted to correspond to the N200 shape and timing (see Figure 14).

The progression of mean voltage for the SI Experimental group across Sessions revealed an initial increase in mean voltage until Session 3 after which it decreased between Session 3 and Session 4 followed by another increase after Session 4 to Session 5 (see Figure 15).

For the effect of Condition and Session on Voltage across the different sessions for the Experimental group in the N200 time-window, another linear mixed model was generated that included SubjectID as random effects (see Figure 16). Data was

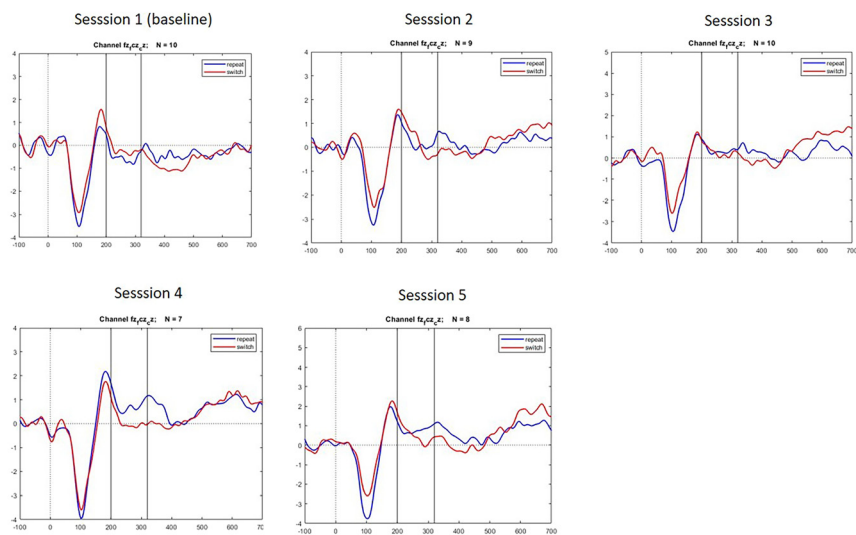


Figure 14: Event-related potentials for the N200 time-window per session.

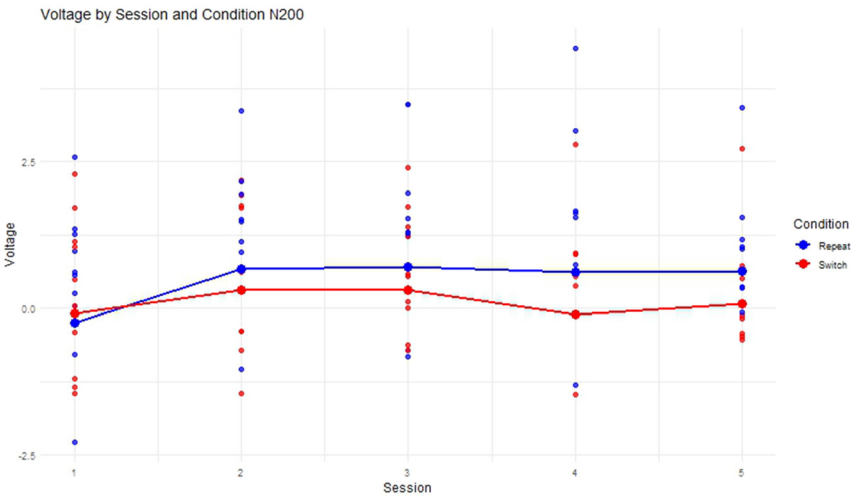


Figure 15: Progression of mean voltage in the switch and repeat condition in the N200 time-window.

normally distributed. None of the effects of Session on Voltage nor the effect of Condition on Voltage were significant (all $p > 0.05$). VIF-values for Session and

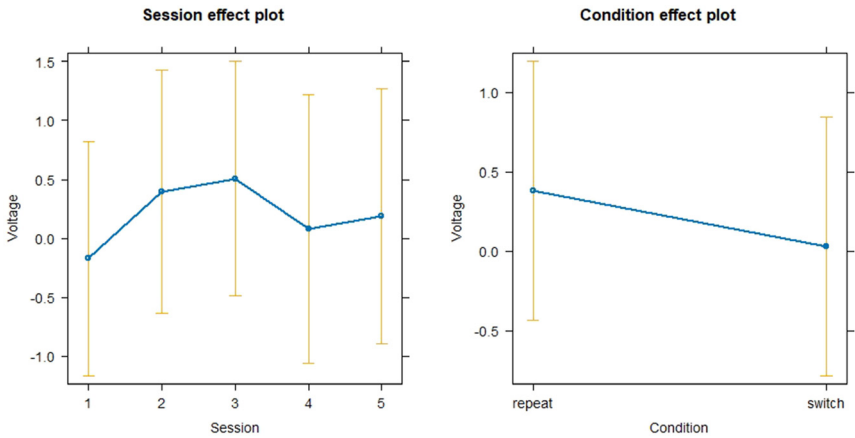


Figure 16: Linear mixed model for the effect of session and condition on voltage in the N200 time-window.

Condition were equal to 1, not indicating any collinearity. Adding an interaction term did not reveal significant interactions as well (all $p > 0.05$).

To see whether BIS is a factor that influences the voltage in the N00 time-window, a linear mixed model with Voltage as dependent variables, Session, Condition and BIC as predictors and Participant as random effect was fitted. Data was normally distributed. Results did not reveal any significant main effects on voltage (all $p > 0.05$). After fitting an interaction model, results did not show any significant two- or three-way interactions.

5 Discussion

The present study aimed to explore how the EF of switching is affected by training language control in the context of SI training by longitudinally testing a group of trainee interpreters in a task switching paradigm.

The results from the TICQ scores indicated an increase in SI competence from the beginning of training to the end of SI training which creates the potential that increased SI competence and therefore increased abilities to deal with the increased processing demands of managing multiple languages while performing overlapping processes of comprehension, translation and speaking positively affects the executive function of task switching.

The lack of significant differences in **reaction times** between the control and Experimental group at baseline level provide no evidence for an initial advantage in

executive functions of the interpreting group compared to the translator group and is in line with the initially formed hypothesis based on findings by Babcock et al. (2017), Dong and Liu (2016) and Rosiers et al. (2019). When comparing the Experimental group and the Control group at baseline level in the **P300 time-window**, the Control group showed significantly lower voltages than the Experimental group which points towards an initial advantage in EF of the interpreters that was not evident in the reaction time data. This contradicts the initially formed hypothesis of no group differences at baseline. However, the interpreters made significantly more mistakes than the translators at baseline testing which indicates that they might have performed at the cost of accuracy. Similar results were found for the **N200 time-window**, where the Control group showed significantly lower voltages than the Experimental group for the N200 time window at baseline testing, suggesting an initial interpreter advantage that is however, challenged by the higher error rate of the interpreters over the translators at baseline testing.

In the **longitudinal testing** of the Experimental group, **behavioral results** revealed that the switch condition showed significantly longer reaction times and decreased accuracy than the repeat condition, which corresponds to the increased processing demands in the switch condition also reported in literature (Declerck et al. 2017; Garbin et al. 2010; Timmer et al. 2017). Both conditions seemed to show steadily decreasing RT and increasing accuracy over the course of testing. In terms of RT, the repeat condition steadily decreased until Session 5 while the switch condition decreased less drastically after Session 4, indicating a ceiling effect that might involve the motor response reaction capability. The results of the linear mixed model did show significant results of all sessions on accuracy, indicating that accuracy increased until the end of testing.

Behavioral switching costs showed an initial increase followed by a decrease peaking at Session 4, again followed by a slight increase. However, a linear mixed model did not show significant results of Session on switching costs. Since the negative peak is extremely low and the following increase only slight, these results might point towards a plateau in RT. The training might have maxed out the motor response to a point where no further increase is possible.

Concerning the longitudinal ERP results in the **P300 time-window**, the switch condition showed significantly smaller voltage than repeat condition, which is in line with the initially formed hypothesis based on the relevant literature that found larger P300 amplitudes with decreasing task demands (Cespón and Carreiras 2020; Gajewski et al. 2018; Jamadar et al. 2015). Figure 12 showed an increase of voltage in the P300 time-window until Session 4 for both conditions, which was followed by a decrease. This is mirrored by the results from the linear mixed model, which found significant effects for Session 3 and Session 4 but not for Session 5, in which voltage seemed to decrease almost back to baseline, as shown in a bridge-shaped curve (see

Figure 13). Taken together with the increasing accuracy and decreasing RT, this might suggest a performance plateau and even an automation of the attentional switching aspect of task switching occurring after Session 4, in line with the hypothesis formed in Section 2.

Regarding the longitudinal ERP-results in the **N200 time-window**, switch trials showed more negative voltage in the N200 time-window than repeat trials which corresponds to the previously formed hypothesis on the basis of relevant ERP literature such as by (Declerck et al. 2017; López Zunini et al. 2019; Timmer et al. 2017). Figure 15 showed that mean voltage increases until Session 3 which is followed by a negative peak in Session 4 and a subsequent increase in voltage for Session 5. The initial increase of mean voltage suggests that participants decreased their allocation of cognitive resources towards interference suppression until Session 3. However, the effects in the linear mixed model for Session were not significant (see Figure 16). Therefore, it is only possible to speak of trends that however do not rely on statistically significant results and should be considered with caution. The drop in voltage at Session 4 could be explained with the Erasmus semester that some of the participants attended after Session 3. Interestingly, the P300 seemed not to be affected by this Erasmus semester which points towards differential executive function mechanisms within task switching, one rather representing interference suppression as indexed by the N200 and one rather representing conflict monitoring and attention switching as indexed by the P300. These might be affected differently by training or the lack thereof. The lack of statistically significant effects of Session on voltage in the N200 time-window might be explained through different approaches: On the one hand, statistical power might have been the problem since a progression of voltage across sessions is visible in Figure 15 but is not reflected in the linear mixed model. On the other hand, participant variability could be the source of the lacking statistical significance. Figure 17 depicts the progression of mean N200 voltage for all participants across sessions. It becomes clear that although inhibitory control peaks at Session 3 for some participants, others show a positive voltage peak at Session 4. This highlights how diverse brain responses to training can be despite having the same training prerequisites. Acknowledging that brains may not react homogeneously to SI training has important implications for SI didactics, calling for a SI curriculum that is more integrated and can consider the needs of the individual students. Going in the same direction, Beaton-Thome (2018) argues for a situated learning approach within a community of practice, suggesting for example student-tutor booth constellations and replacing the final exam at the end of the SI training with a more continuous assessment of SI competence (Beaton-Thome 2018: 158–160).

Lastly, it is also possible that inhibitory control is just not as implicated in the executive function of switching in the form it is applied to cope with the enhanced processing demands of controlling two languages during simultaneous interpreting.

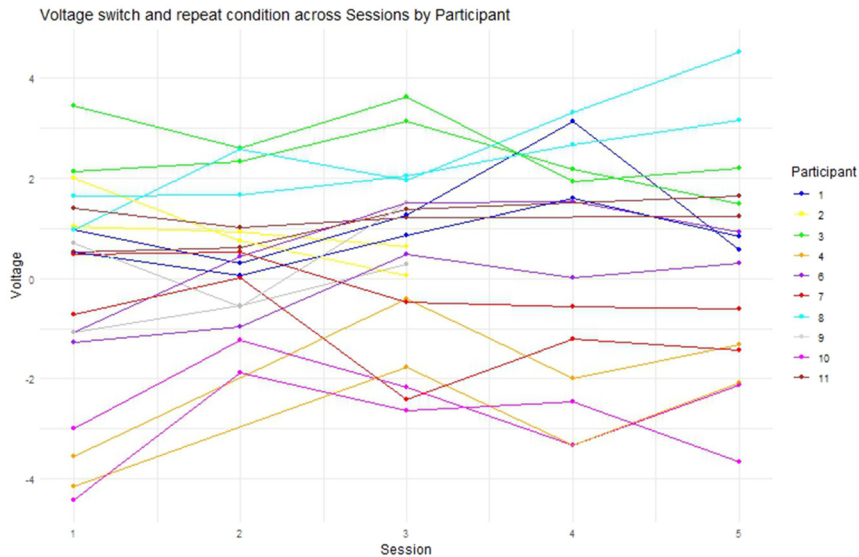


Figure 17: Mean voltage per condition, session, and participant in the N200 time-window.

In comparison to switching and working memory updating, inhibitory control or rather interference suppression is the executive functions that has been shown to be affected by an interpreter advantage the least (see Nour et al. 2020 for a review).

6 Conclusions

To sum up, the most critical finding from the present study is that simultaneous interpreting training seems to affect the executive function of task switching, however, more specifically, the present study has shown that SI training specifically seems to affect the aspect of allocating attentional resources to switching between tasks but not suppressing interference during task switching. This is in line with previous findings, showing improved RT and accuracy after SI training. Participants in the interpreting group therefore seemed to improve their ability to switch attention as their SI training progressed and therefore their SI competence increased, improving their skills to accommodate the increased processing demands to manage multiple languages while performing overlapping processes of language comprehension, translation, and speech production. Furthermore, as mapped in models of general skill acquisition (see Chein and Schneider 2012) as well as models from bilingualism (see Paap 2018) and studies on SI training and executive functions

(see Dong and Liu 2016), the improvement in attention switching seemed to reach a plateau near the end of SI training when SI competence had been built to a certain extent. Especially the combination of electrophysiological measures and behavioral measures in the present study was quite insightful as it provided details into the reaction of the brain that does not contain a time lag for the motor response as is the case in purely behavioral studies. Furthermore, applying a methodology that went beyond pre- and post tests, gathering data continuously throughout SI training was crucial to reveal performance peaks that would otherwise have been disguised.

Although these findings need to be replicated with neuroimaging methods to validate this for certain, the results highlight the plastic nature of the brain in reaction to experience in managing multiple languages in one brain in the circumstance of extreme language control that is simultaneous interpreting.

Overall, this study contributes to shedding more light on the temporal progression of the executive function of switching during interpreting training and highlights the importance of neurolinguistic research methods in translation and interpreting studies.

Future studies should investigate whether translation training affects executive functions, which should involve longitudinal designs using electrophysiological measures. Also, it could be interesting to investigate the short-term effects of translation and interpreting training. Studies on language learning have found microstructural changes in the cortex after less than 1 h of new word language learning (see Hofstetter et al. 2017). Therefore, exposing participants to a translation, interpreting or post-editing task and administering an executive function task afterwards, coupled with a control condition, could provide intriguing results.

7 Limitations

The aforementioned between-participant variability is strongly connected to the limitations of this study. The number of participants is very small, which, although quite common in longitudinal neuroscientific studies, negatively impacts statistical power and fails to balance out differences in participant performance. This is an important point since great voltage variation per participant was visible in the single sessions highlighting that the individual's brains respond quite differently to the same kind of training. Furthermore, the Control group was tested only once to compare differences between groups at baseline level. As mentioned in the methodology, this was due to the necessary Covid-19 restrictions that made lab use involving participants difficult. Ideally, a Control group would have been selected simultaneously to the SI group and would have been tested throughout their translation training parallel to the interpreting group.

References

- Abutalebi, Jubin, Pasquale Anthony Della Rosa, Anna Kaarina Castro Gonzaga, Roland Keim, Albert Costa & Daniela Perani. 2013. The role of the left putamen in multilingual language production. *Brain and Language* 125(3). 307–315.
- Abutalebi, Jubin & David Green. 2007. Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics* 20(3). 242–275.
- Abutalebi, Jubin & David Green. 2008. Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes* 23(4). 557–582.
- Abutalebi, Jubin & David Green. 2016. Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition* 19(4). 689–698.
- Babcock, Laura, Mariagrazia Capizzi, Sandra Arbula & Antonino Vallesi. 2017. Short-term memory improvement after simultaneous interpretation training. *Journal of Cognitive Enhancement* 1(3). 254–267.
- Babcock, Laura & Antonino Vallesi. 2017. Are simultaneous interpreters expert bilinguals, unique bilinguals, or both? *Bilingualism: Language and Cognition* 20(2). 403–417.
- Baene, Wouter de, Wouter Duyck, Marcel Brass & Manuel Carreiras. 2015. Brain circuit for cognitive control is shared by task and language switching. *Journal of Cognitive Neuroscience* 27(9). 1752–1765.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). <https://doi.org/10.18637/jss.v067.i01>.
- Beaton-Thome, Morven. 2018. Situated expertise in interpreting. In Barbara Ahrens, Silvia Hansen-Schirra, Monika Krein-Mühle, Michael Schreiber & Ursula Wienen (eds.), *Translation – Didaktik – Kompetenz*, 145–167. Berlin: Franke & Timme GmbH.
- Becker, Maxi, Torsten Schubert, Tilo Strobach, Jürgen Gallinat & Simone Kühn. 2016. Simultaneous interpreters vs. professional multilingual controls: Group differences in cognitive control as well as brain structure and function. *NeuroImage* 134. 250–260.
- Bialystok, Ellen, Fergus Craik & Gigi Luk. 2008. Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 34(4). 859–873.
- Braver, Todd, Jeremy Reynolds & David I. Donaldson. 2003. Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron* 39(4). 713–726.
- Bunge, Silvia, Nicole Dudukovic, Moriah Thomason, Chandan Vaidya & John Gabrieli. 2002. Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron* 33(2). 301–311.
- Calabria, Marco, Albert Costa, David Green & Jubin Abutalebi. 2018. Neural basis of bilingual language control. *Annals of the New York Academy of Sciences* 1426(1). 221–235.
- Cespón, Jesús. 2021. Neural processing underlying executive functions in bilinguals: “Heads I Win, Tails You Lose”. *Frontiers in Human Neuroscience* 15. 710905.
- Cespón, Jesús & Manuel Carreiras. 2020. Is there electrophysiological evidence for a bilingual advantage in neural processes related to executive functions? *Neuroscience and Biobehavioral Reviews* 118. 315–330.
- Chein, Jason & Walter Schneider. 2012. The brain’s learning and control architecture. *Current Directions in Psychological Science* 21(2). 78–84.
- Chen, Yun, Bihua Cao, Liufang Xie, Jianxiao Wu & Fuhong Li. 2022. Proactive and reactive control differ between task switching and response rule switching: Event-related potential evidence. *Neuropsychologia* 172. 108272.

- Christoffels, Ingrid, Annette Degroot & Judith Kroll. 2006. Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language* 54(3). 324–345.
- Cohen, Jacob. 1977. *Statistical power analysis for the behavioral sciences*. New York, NY: Acad. Press.
- Costa, Albert, Mireia Hernández, Jordi Costa-Faidella & Núria Sebastián-Gallés. 2009. On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition* 113(2). 135–149.
- Costa, Albert, Mireia Hernández & Núria Sebastián-Gallés. 2008. Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition* 106(1). 59–86.
- Declerck, Mathieu, Jonathan Grainger, Iring Koch & Andrea M. Philipp. 2017. Is language control just a form of executive control? Evidence for overlapping processes in language switching and task switching. *Journal of Memory and Language* 95. 138–145.
- Declerck, Mathieu, Gabriela Meade, Katherine Midgley, Phillip Holcomb, Ardi Roelofs & Karen Emmorey. 2021. On the connection between language control and executive control—an ERP study. *Neurobiology of Language (Cambridge, Mass.)* 2(4). 628–646.
- Delorme, Arnaud & Scott Makeig. 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* 134(1). 9–21.
- Diamond, Adele. 2013. Executive functions. *Annual Review of Psychology* 64. 135–168.
- Dong, Yanping & Yuhua Liu. 2016. Classes in translating and interpreting produce differential gains in switching and updating. *Frontiers in Psychology* 7. 1297.
- Dong, Yanping & Fei Zhong. 2017. Interpreting experience enhances early attentional processing, conflict monitoring and interference suppression along the time course of processing. *Neuropsychologia* 95. 193–203.
- Elmer, Stefan, Jürgen Hänggi & Lutz Jäncke. 2014. Processing demands upon cognitive, linguistic, and articulatory functions promote grey matter plasticity in the adult multilingual brain: Insights from simultaneous interpreters. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior* 54. 179–189.
- Elmer, Stefan, Jürgen Hänggi, Martin Meyer & Lutz Jäncke. 2011. Differential language expertise related to white matter architecture in regions subserving sensory-motor coupling, articulation, and interhemispheric transfer. *Human Brain Mapping* 32(12). 2064–2074.
- Elmer, Stefan, Martin Meyer & Lutz Jancke. 2010. Simultaneous interpreters as a model for neuronal adaptation in the domain of language processing. *Brain Research* 1317. 147–156.
- Folstein, Jonathan R. & Cyma van Petten. 2008. Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology* 45(1). 152–170.
- Gajewski, Patrick, Nicola Ferdinand, Jutta Kray & Michael Falkenstein. 2018. Understanding sources of adult age differences in task switching: Evidence from behavioral and ERP studies. *Neuroscience and Biobehavioral Reviews* 92. 255–275.
- Garbin, Gabriele, Ana Sanjuan, Cristina Forn, Juan Carlos Bustamante, Aina Rodriguez-Pujadas, Vicente Belloch, Mireia Hernandez, Albert Costa & César Avila. 2010. Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. *NeuroImage* 53(4). 1272–1278.
- Grant, David & Berg Esta. 1948. *PsycTESTS Dataset (PsycTESTS Dataset)*. American Psychological Association (APA).
- Grogan, Alice, Oiwi Parker Jones, Nilufa Ali, Jenny Crinion, S. Orabona, Marie Luise Mechias, Sue Ramsden, David Green & Cathy Price. 2012. Structural correlates for lexical efficiency and number of languages in non-native speakers of English. *Neuropsychologia* 50(7). 1347–1352.
- Grundy, John, John Anderson & Ellen Bialystok. 2017. Neural correlates of cognitive processing in monolinguals and bilinguals. *Annals of the New York Academy of Sciences* 1396(1). 183–201.

- Henrard, Sébastien & Agnès van Daele. 2017. Different bilingual experiences might modulate executive tasks advantages: Comparative analysis between monolinguals, translators, and interpreters. *Frontiers in Psychology* 8. 1870.
- Hernandez, Arturo, Antígona Martínez & Kathryn Kohnert. 2000. In search of the language switch: An fMRI study of picture naming in Spanish-English bilinguals. *Brain and Language* 73(3). 421–431.
- Hervais-Adelman, Alexis & Laura Babcock. 2020. The neurobiology of simultaneous interpreting: Where extreme language control and cognitive control intersect. *Bilingualism (Bilingualism: Language and Cognition)* 23(4). 740–751.
- Hervais-Adelman, Alexis, Barbara Moser-Mercer & Narly Golestani. 2015. Brain functional plasticity associated with the emergence of expertise in extreme language control. *NeuroImage* 114. 264–274.
- Hofstetter, Shir, Naama Friedmann & Yaniv Assaf. 2017. Rapid language-related plasticity: Microstructural changes in the cortex after a short session of new word learning. *Brain Structure & Function* 222(3). 1231–1241.
- Hosoda, Chihiro, Kanji Tanaka, Tadashi Nariai, Manabu Honda & Takashi Hanakawa. 2013. Dynamic neural network reorganization associated with second language vocabulary acquisition: A multimodal imaging study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 33(34). 13663–13672.
- Jacob, Gunnar, Moritz Jonas Schaeffer, Katharina Oster & Silvia Hansen-Schirra. 2024. The psycholinguistics of shining-through effects in translation: Cross-linguistic structural priming or serial lexical co-activation? *Applied Psycholinguistics (Applied Psycholinguistics)* 45(3). 542–566.
- Jamadar, Sharna D., Renate Thienel & Frini Karayanidis. 2015. Task switching processes. In Arthur W. Toga (ed.), *Brain mapping: An Encyclopedic reference*, 327–335. Waltham: Academic Press.
- Jodo, Eiichi & Yukihiko Kayama. 1992. Relation of a negative ERP component to response inhibition in a Go/No-go task. *Electroencephalography and Clinical Neurophysiology* 82(6). 477–482.
- Kaiser, Anelis, Leila Eppenberger, Renata Smieskova, Stefan Borgwardt, Esther Kuenzli, Ernst-Wilhelm Radue, Cordula Nitsch & Kerstin Bendfeldt. 2015. Age of second language acquisition in multilinguals has an impact on gray matter volume in language-associated brain areas. *Frontiers in Psychology* 6. 638.
- Khateb, Asaid, Jubin Abutalebi, Christoph Michel, Alan Pegna, Hannelore Lee-Jahnke & Jean-Marie Annoni. 2007. Language selection in bilinguals: A spatio-temporal analysis of electric brain activity. *International Journal of Psychophysiology* 65(3). 201–213.
- Klein, Carina, Silvana Metz, Stefan Elmer & Lutz Jäncke. 2018. The interpreter's brain during rest – hyperconnectivity in the frontal lobe. *PLoS One* 13(8). e0202600.
- Kopp, Bruno, Fred Rist & Uwe Mattler. 1996. N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology* 33(3). 282–294.
- Lisefeld, Heinrich René & Markus Janczyk. 2019. Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods* 51(1). 40–60.
- López Zunini, Rocío, Cassandra Morrison, Shanna Kousaie & Vanessa Taler. 2019. Task switching and bilingualism in young and older adults: A behavioral and electrophysiological investigation. *Neuropsychologia* 133. 107186.
- Lövdén, Martin, Elisabeth Wenger, Johan Mårtensson, Ulman Lindenberger & Lars Bäckman. 2013. Structural brain plasticity in adult learning and development. *Neuroscience and Biobehavioral Reviews* 37(9 Pt B). 2296–2310.
- Luk, Gigi, David Green, Jubin Abutalebi & Cheryl Grady. 2011. Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes* 27(10). 1479–1488.

- Marin-Marín, Lidón, Víctor Costumero, César Ávila & Christos Pliatsikas. 2022. Dynamic effects of immersive bilingualism on cortical and subcortical grey matter volumes. *Frontiers in Psychology* 13. 886222.
- Mechelli, Andrea, Jenny Crinion, Uta Noppeney, John O'Doherty, John Ashburner, Richard S. Frackowiak & Cathy J. Price. 2004. Neurolinguistics: Structural plasticity in the bilingual brain. *Nature* 431(7010). 757.
- Meiran, Nachshon. 1996. Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 22(6). 1423–1442.
- Miyake, Akira & Naomi Friedman. 2012. The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science* 21(1). 8–14.
- Miyake, Akira, Naomi Friedman, Michael Emerson, Alexander Witzki, Amy Howerter & Tor Wager. 2000. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology* 41(1). 49–100.
- Monsell, Stephen. 2003. Task switching. *Trends in Cognitive Sciences* 7(3). 134–140.
- Nour, Soudabeh, Esli Struys, Evy Woumans, Ily Hollebeke & Hélène Stengers. 2020. An interpreter advantage in executive functions? *Interpreting* 22(2). 163–186.
- Paap, Kenneth. 2018. Bilingualism in cognitive science. In Annick de Houwer & Lourdes Ortega (eds.), *The Cambridge handbook of bilingualism*, 435–465. University Erfurt, Germany; Georgetown University, Washington DC: Cambridge University Press.
- Paap, Kenneth & Zachary Greenberg. 2013. There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology* 66(2). 232–258.
- Peirce, Jonathan. 2007. PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods* 162(1–2). 8–13.
- Periáñez, Jose & Francisco Barceló. 2009. Updating sensory versus task representations during task-switching: Insights from cognitive brain potentials in humans. *Neuropsychologia* 47(4). 1160–1172.
- Pliatsikas, Christos. 2020. Understanding structural plasticity in the bilingual brain: The Dynamic Restructuring Model. *Bilingualism: Language and Cognition* 23(2). 459–471.
- Pliatsikas, Christos, Vincent DeLuca, Elisavet Moschopoulou & James Saddy. 2017. Immersive bilingualism reshapes the core of the brain. *Brain Structure & Function* 222(4). 1785–1795.
- Poarch, Gregory & Ellen Bialystok. 2015. Bilingualism as a model for multitasking. *Developmental Review* 35. 113–124.
- Polich, John. 2007. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* 118(10). 2128–2148.
- Posit team. 2024. *RStudio: Integrated Development Environment for R: Posit Software, PBC*. Boston, MA.
- Price, Cathy, David Green & Roswitha von Studnitz. 1999. A functional imaging study of translation and language switching. *Brain: A Journal of Neurology* 122(Pt 12). 2221–2235.
- Prior, Anat & Tamar Gollan. 2011. Good language-switchers are good task-switchers: Evidence from Spanish-English and Mandarin-English bilinguals. *Journal of the International Neuropsychological Society* 17(4). 682–691.
- Rosiers, Alexandra, Evy Woumans, Wouter Duyck & June Eyckmans. 2019. Investigating the presumed cognitive advantage of aspiring interpreters. *Interpreting. International Journal of Research and Practice in Interpreting* 21(1). 115–134.
- Rossi, Eleonora, Hu Cheng, Judith Kroll, Michele Diaz & Sharlene Newman. 2017. Changes in white-matter connectivity in late second language learners: Evidence from diffusion tensor imaging. *Frontiers in Psychology* 8(2040). 1–15.

- Schaeffer, Moritz, David Huepe, Silvia Hansen-Schirra, Sascha Hofmann, Edinson Muñoz, Boris Kogan, Eduar Herrera, Agustín Ibáñez & Adolfo García. 2020. The translation and interpreting competence questionnaire: An online tool for research on translators and interpreters. *Perspectives* 28(1). 90–108.
- Selya, Arielle, Jennifer Rose, Lisa Dierker, Donald Hedeker & Robin Mermelstein. 2012. A practical guide to calculating Cohen's $f(2)$, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology* 3. 111.
- Timmer, Kalinka, John Grundy & Ellen Bialystok. 2017. Earlier and more distributed neural networks for bilinguals than monolinguals during switching. *Neuropsychologia* 106. 245–260.
- van de Putte, Eowyn, Wouter de Baene, Lorna García-Pentón, Evy Woumans, Aster Dijkgraaf & Wouter Duyck. 2018. Anatomical and functional changes in the brain after simultaneous interpreting training: A longitudinal study. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior* 99. 243–257.
- van den Noort, Maurits, Esli Struys, Peggy Bosch, Lars Jaswetz, Benoît Perriard, Sujung Yeo, Pia Barisch, Katrien Vermeire, Sook-Hyun Lee & Sabina Lim. 2019. Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences* 9(3). 27.
- van der Linden, Lize, Eowyn van de Putte, Evy Woumans, Wouter Duyck & Arnaud Szmalec. 2018. Does extreme language control training improve cognitive control? A comparison of professional interpreters, L2 teachers and monolinguals. *Frontiers in Psychology* 9. 1998.
- Verreyt, Nele, Evy Woumans, Davy Vandelandotte, Arnaud Szmalec & Wouter Duyck. 2016. The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition* 19(1). 181–190.
- Weissberger, Gali, Tamar Gollan, Mark Bondi, Lindsay Clark & Christina Wierenga. 2015. Language and task switching in the bilingual brain: Bilinguals are staying, not switching, experts. *Neuropsychologia* 66. 193–203.
- Yudes, Carolina, Pedro Macizo & Teresa Bajo. 2011. The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in Psychology* 2. 309.
- Zuur, Alain, Elena Ieno & Chris Elphick. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1(1). 3–14.