“We never use the same brain twice” – How does interpreting training affect plasticity in domain-general cognitive control? A longitudinal EEG-study

Inauguraldissertation

zur Erlangung des Akademischen Grades eines

Dr. phil.

vorgelegt dem Fachbereich Translations-, Sprach- und Kulturwissenschaft (06) der
Johannes-Gutenberg-Universität Mainz

von

Ann-Kathrin Habig

aus Fulda

Referent/in: Prof. Dr. Silvia Hansen-Schirra

1. Korreferent/in: Prof. Dr. Alexis Hervais-Adelmann

2. Korreferent/in: Prof. Dr. Morven Beaton-Thome

Tag des Prüfungskolloquiums: 29.09.2025

Abstract

The aim of this dissertation is to examine how simultaneous interpreting training as a context where extreme language control is needed affects the three main executive functions of inhibition, switching and working memory updating. To this end, this dissertation examines the basis of brain plasticity and learning in a linguistic context, defines executive functions and their role in bilingual language processing and highlights simultaneous interpreting as a modality of bilingualism that clearly taxes both language control as well as executive functions. By emphasizing how interpreters use interpreting-specific competences to deal with the increased control demands and depicting the processes of interpreting competence acquisition in interpreting training, a parallel to general skill acquisition processes from the cognitive sciences is drawn. To conclude the theoretical background, relevant studies from cognitive bilingualism studies as well as cognitive translation and interpreting studies on the impact of simultaneous interpreting on executive functions are presented.

Based on these theoretical underpinnings, the methodology and results of a longitudinal event-related potential study are presented. The performance in three executive function tasks of a group of interpreting students was compared to translation students at baseline before their respective Master studies. The performance of the interpreting group was then tested subsequently at the end of every semester for four semesters until the end of their Master studies of Conference Interpreting while event-related potential data was also collected. The event-related potential data was triangulated with behavioural data as well as competence data. The rare combination of a longitudinal design that goes beyond pre- and post-tests and was paired with the neuroscientific method of electroencephalography allows for a much more fine-grained look at how executive functions are affected by the process of learning how to interpret simultaneously.

Results are presented as discussed in relation to previous research. The findings not only enrich the research landscape and knowledge base surrounding executive functions and language control by offering unique insights into temporal aspects of neural processes but also provide implications for the didactics of interpreting training.

Finally, room for enhancement in the conducted study as well as opportunities for future research are discussed.

Content

List of Figures

List of Tables

Abbreviations

ACC	anterior cingulate cortex
ANT	attention network task
AoA	Age of Acquisition
AOI	area of interest
ASL	American Sign Language
BA	Bachelor
BAPSS	bilingual anterior-to-posterior and subcortical shift
BEPA	bilingual executive processing advantage
BIA	Bilingual Interactive Activation
BICA	bilingual inhibitory control advantage
BL	bilingual
CC	corpus callosum
CSM	colour-shape mix
DLPFC	dorsolateral prefrontal cortex
EEG	electroencephalography
EF	executive function
ERD	event-related desynchronization
ERP	event-related potential
ERS	event-related synchronization
ERSP	event-related spectral perturbation
ESIT	École Supérieure d'Interprêtes et Traducteures
FTSK	Fachbereich Translations-, Sprach- und Kulturwissenschaft
GAM	Generated Abstract Memory
ICA	independent component analysis
ICM	Inhibitory Control Model
IFC	inferior frontal cortex
IFG	inferior frontal gyrus

IFOF	inferior fronto-occipital fasciculus
ILF	inferior longitudinal fasciculus
fNIRS	functional near-infrared spectroscopy
IPC	inferior parietal cortex
IPL	inferior parietal lobe
LEAP	Language and Proficiency Questionnaire
LKM	Leipziger Kompetenzmodell der Dolmetschdidaktik
LRP	lateralized readiness potential
MA	Master
MDA	multigroup discriminant function analysis
MFG	middle frontal gyrus
MRI	magnetic resonance imaging
MTG	middle temporal gyrus
NTM	neuroarchitectural translation model
PCA	principal component analysis
PCG	precentral gyrus
PFC	prefrontal cortex
PET	positron emission tomography
PL	parietal lobe
RT	reaction time
SAS	supervisory attentional system
SCE	sequential congruency effect
SD	standard deviation
SE	standard error
SFG	superior frontal gyrus
SH	shadowing
SI	simultaneous interpreting
SL	source language
SLF	superior longitudinal fasciculus

SMA	supplementary motor area
SMG	supramarginal gyrus
SRC	stimulus-response-conflict
ST	source text
STG	superior temporal gyrus
STM	short-term memory
STS	superior temporal sulcus
TICQ	Translation and Interpreting Competence Questionnaire
TL	Target Language
TPR	translation process research
VIF	variance inflation factor
VMPFC	ventromedial prefrontal cortex
VLPFC	ventrolateral prefrontal cortex
WCST	Wisconsin Card Sorting Test
WM	working memory

1. Introduction

"Our brains are adaptable; indeed, we never use the same brain twice"

(Turner 2019)

Brain plasticity has been considered the basis of learning and, although "never using the same brain twice" is clearly hyperbole, our brains do have the ability to change dynamically and adapt to external influences, ultimately being shaped by the lives we lead (see Costandi 2016: 2). Communicating through language, be it spoken language or other modalities, is an integral part of our human existence and therefore can be considered as one of the most crucial external influences our brain is subject to. Since bilingualism, or, managing more than one language in the brain, has been observed to result in structural and functional changes and in some cases even in cognitive benefits that lie outside the linguistic domain (see Mechelli et al. (2004), Grogan et al. (2012), Rossi et al. (2017), Berken et al. (2016), Bialystok et al. (2004; 2005)), the search for the circumstances and mechanisms under which these cognitive changes and benefits occur has been the focus of many researchers (see e.g. Abutalebi & Green (2008), Bialystok et al. (2005), Paap & Greenberg (2013), Verreyt et al. (2016), Hervais-Adelman et al. (2015a), Elmer et al. (2011)). The basis of these research questions is that in bilinguals, both of their languages are active to a certain degree, resulting in increased control demands to manage both languages and avoid unwanted interferences (see e.g. Dijkstra & van Heuven (1998), Green (1998), Meuter & Allport (1999), Paradis (1994)). Furthermore, neural regions associated with both linguistic processing and well as domain-general executive control partly overlap neurally, which has highlighted so called executive functions (EF) as an actively contributing mechanism for the effective management of languages during bilingual language processing (see e.g. Abutalebi & Green (2007), (2008), (2013), Hervais-Adelman & Babcock (2020), Baene et al. (2015), Garbin et al. (2010)).

Simultaneous interpreting (SI) has for some time now become the centre of interest of researchers investigating how the multilingual brain changes in response to different cognitive control demands. In addition to the acquisition of interpreting-specific skills to manage the high processing demands of the interpreting task at hand, interpreters need to keep both of their languages at a high level of activation during simultaneous interpreting while also avoiding unwanted interference from both languages. Overlapping processes of listening, word identification, translation and speech production (see Frauenfelder & Schriefers 1997) create higher processing demands than other bilingual speech contexts, which in turn requires a higher degree of cognitive control to coordinate the various brain networks carrying out these different tasks and therefore taxes executive functions to a greater extent than other bilingual interactional contexts (see García (2014), Christoffels et al. (2006), Hervais-Adelman & Babcock (2020), Yudes et al. (2011), Henrard & van Daele (2017), van der Linden et al. (2018)). Consequently, neurocognitive models of SI that are based on empirical research findings have been established (see e.g. Fabbro (1999), Garcia (2019), Hervais-Adelman & Babcock (2020)).

These insights have contributed to the interpreter advantage research that investigates executive functions of simultaneous interpreters in direct comparison with language groups characterized by different bilingual modalities such as translators, language teachers or general bilinguals without interpreting competence (see e.g. Yudes et al. (2011), Dong & Liu (2016), Babcock et al. (2017), Zhong & Dong (2024), van de Putte et al. (2018)). As in the bilingual advantage hypothesis, results are quite heterogenous, often depending on the research method that is applied (neuroscientific or behavioural) and the executive function that is researched (see e.g. Henrard & van Daele (2017) but see also e.g. van der Linden et al. (2018)). However, meta-analyses have identified that the executive functions of task-switching and working memory (WM) updating play a primary role during SI (see Nour et al. (2020) and Hu & Fan (2021)).

The question whether or not there is an improvement in executive function as an effect of SI, however, only targets participant groups as a snapshot in time and does not address another basal and important aspect in this research field: How and when do advantages in executive functions as effects of SI come to pass? Since SI is an acquired skill that follows skill acquisition patterns based in general cognitive science, simultaneous interpreting training is especially relevant in this context. Interpreting students undergo a learning process of acquiring interpreting specific skills and refining language skills (see Kutz (2010), Kalina (2006), Albl-Mikasa (2012), Riccardi (2005), Hoffman (1997), Moser-Mercer et al. (2000)) while learning to interpret simultaneously which requires differential involvement of different executive functions. However, this skill acquisition process, how executive functions are affected by the process of learning how to interpret simultaneously, has not been given much attention in SI research.

Hence, I conducted a longitudinal study using both neuroscientific methods, behavioural data as well as competence data to investigate how executive functions change during SI training as a function of increased cognitive control demands and interpreting-specific skill acquisition. My aim was to specifically monitor the progression of three different executive functions during SI training to gain insights into the black box of simultaneous interpreter's brains that go beyond pre- and post-test designs.

Simultaneous interpreting as a research subject is situated in the field of translation and interpreting studies. However, the methodology applied in this dissertation heavily draws on cognitive bilingualism research. Both interpreting studies as well as bilingualism research have translation as a common denominator. Nonetheless, they approach translation from two different angles. Historically, translation was used as a task in bilingualism research to study and understand the mechanisms of language processing in the brain. For example, Kroll & Stewart (1994) applied a translation task, the results of which lead to the conceptualization of the Revised Hierarchical Model (RHM), one of the most influential models of bilingual language processing. Furthermore, early case studies of bilingual aphasic patients applied translation tasks to show that language processing is disrupted (see e.g. Paradis et al. (1982)). In interpreting studies however, translation in the mode of simultaneous interpreting is an acquired skill and a profession. In the past, cognitive translation studies have been criticized for not relying on insights from cognitive bilingualism research in different aspects (see e.g. García (2015)), despite the fact that translation requires controlling information in the multilingual mental lexicon, an aspect of cognition that psycholinguists have spent decades researching (García 2015: 10–11). Furthermore, although interpreting studies have drawn on cognitive models of language processing from bilingualism research in recent years to establish neurocognitive models of interpreting (see Garcia (2019) and Hervais-Adelman & Babcock (2020)), bilingualism research does not integrate findings from cognitive interpreting studies to that extent, although it could surely benefit from it. This dissertation attemps to bridge this gap by taking the fundamentals of bilingual language processing as its basis and then applying them to the setting of simultaneous interpreting, using well-established methodology from cognitive bilingualism research and finally contextualizing the results as contributions to the fields of cognitive interpreting studies and bilingualism research.

1.1 Research objective

While there is a substantial body of research investigating the interpreting advantage (see e.g. Hu & Fan (2021) or García et al. (2020) for reviews), there are only a handful of studies that investigate how SI training affects executive functions (e.g. Dong & Zhong (2017)). Although cross-sectional studies have contributed to a better understanding of the involvement of executive functions during SI, only few studies have investigated how SI training affects executive functions in a longitudinal design (see Dong & Liu (2016), Babcock et al. (2017) and van de Putte et al. (2018)) which is, however, important to not merely consider a snapshot in time. Moreover, studies that have applied a longitudinal design, did so by using pre- and post- designs, usually testing groups of participants at the beginning and at the end of their respective training. While pre- and post-test designs can potentially give insights into whether there is an improvement or change in executive functions from the start of training to the end of training, they cannot provide information on what happens in-between these testing points. Assuming executive functions progress in alignment with interpreting skill acquisition and might therefore be characterized by peaks, plateaus or a renormalization of the progress, pre- and post-test designs do not suffice to depict this development. Furthermore, the studies that are investigating the effect of SI training on executive functions rarely apply neuroscientific measures that have, however, the potential to provide important details about cognitive processes in a granularity that behavioural measures cannot.

Hence, I aim to answer the research question of how SI training affects the three executive functions of inhibition, switching and working-memory updating by conducting a longitudinal study that goes beyond pre- and post-test measures and also takes skill acquisition models from cognitive science as well as interpreting studies into consideration.

With the aim of monitoring the changes in the reaction of the brain to certain executive function tasks, I decided to use electroencephalography (EEG), more specifically event-related potentials (ERP), in triangulation with behavioural data (reaction time and accuracy data) and self-rated as well as objective competence data. The EEG and especially ERPs are characterized by a very high temporal resolution, making it possible to precisely observe different brain reactions over time (see e.g. Luck (2005)). In contrast to reaction time measures that also contain the time it takes to give a motor response, ERPs are a more direct way to cognitive processes. Applying linear mixed effects models that allow the inclusion of various independent variables as well as random effect to account for participant variability, makes it possible to implement statistical rigour and obtain reliable effects.

In addition to the question whether executive functions progress or improve at all, I am especially interested in whether executive functions show a plateau or even a decrease that is accompanied by a stable behavioural performance, in accordance with neuroscientific and neurolinguistic models of skill acquisition and language learning (see e.g. Lövdén et al. (2013) and Pliatsikas (2020)).

Results contribute to the already existing research on SI and executive functions, complement the findings from previous studies on the effect of SI training on executive functions and enrich the existing literature on executive functions in cognitive bilingual studies. Showing that executive functions do not progress continuously throughout training but might be subject to performance plateaus or renormalization after performance peak, or that the progression of executive functions might be affected by pauses of training such as semesters abroad, might have various implications. On the one hand, affirming results would contribute to the understanding of how SI as a form of extreme language control affects the plasticity of bilingual brains. On the other hand, it can have implications for didactics: if performance and electrophysiological markers decrease after a semester abroad where no interpreting training was followed, it might be reasonable to offer additional preparatory courses before resuming interpreting training again after the semester abroad or to supply the SI trainees with material they can use to train during their semester abroad. Expanding the implications of my results to didactics ultimately benefits the simultaneous interpreting trainees, since the insights about the progress of cognitive control functions can lead to the Master's programme in Conference Interpreting being better tailored to the needs of the students.

1.2. Chapter outline

Chapter 2 introduces how bilingualism as well as SI can bring about changes in the brain by initially defining the term plasticity and learning-induced plasticity in a language-independent manner. This chapter lays the foundation for explaining how the brain changes in reaction to external influences, highlights the basic mechanisms of learning and skill acquisition by introducing relevant models from cognitive science and describes how expertise can be transferred from one domain to another if the underlying skills share common cognitive processes, such as executive functions. Chapter 3 starts out by defining bilingualism and explaining the caveats of categorizing bilingual individuals. Building on the knowledge of neural plasticity from the previous chapter, Chapter 3 presents neurolinguistics studies that investigate how bilingualism affects the brain functionally as well as structurally and continues to provide evidence for linguistically-induced brain plasticity from cognitive translation and interpreting studies. Specifically, brain regions that are affected by language-induced brain plasticity are highlighted, marking the first mention of overlapping involvement of domain-general executive functions in linguistic processes.

Chapter 4 is dedicated to defining executive functions. First of all, the term executive function is defined in general, followed by a detailed description of the three main executive functions inhibition, switching (shifting) and working-memory updating. In a dedicated sub-chapter, these executive functions are further differentiated and mapped onto tasks that make it possible to test the respective executive function. Chapter 5 marks the point where the aforementioned isolated knowledge about executive functions and linguistic processes converges. It begins by posing a key question in cognitive bilingualism research: how do bilinguals manage more than one language in their mind without being subject to unwanted interference from one language when another is used? To elucidate on this question, relevant psycholinguistic models of both word identification as well as word production are delineated in separate subchapters. By doing so, first mentions of the involvement of executive functions in these processes are highlighted. Building on this, another subchapter is dedicated to investigating the neural overlap of domain-general executive functions and language-specific processes during language processing. Relevant studies investigating this overlap with behavioural as well as neuroscientific methods are presented, and a language control network depicting the involved brain regions implicated in executive functions and language processing is illustrated. In the final subchapter of Chapter 5, the bilingual advantage hypothesis is presented, integrating knowledge about language-induced brain plasticity and skill acquisition from Chapter 2 and Chapter 3 as well as knowledge about the involvement of executive functions in bilingual language processing from Chapter 4 and the current Chapter 5. The chapter concludes with presenting factors of the bilingual experience that may modulate the change in executive functions that constitutes the basis of the bilingual advantage hypothesis, putting simultaneous interpreting in the focus as an instance of extreme language control.

In order to describe SI as a cognitive task, Chapter 6 starts out by defining SI and setting it apart from other forms of translation and modalities of bilingualism. Key literature and associated models that highlight important SI-specific processes are presented and the involvement of executive functions in these models and processes is emphasized. Subsequently, interpreting competence models are presented and SI competence acquisition models are accentuated to draw a parallel to skill acquisition models from cognitive science mentioned in Chapter 2. Pertinent studies using both behavioural as well as neuroscientific methods are presented in another subchapter that indicates how SI as an extreme form of language control affects the brain structurally and functionally, complementing studies from Chapter 3 that concern language-induced brain plasticity in general. Finally, studies using both behavioural as well as different neuroscientific methods are presented that investigate how SI and SI training affect executive functions.

Chapter 7 provides an interim summary of all previous chapters, highlighting the key points that are relevant to the research question of this dissertation. Furthermore, Chapter 7 presents the research gap as well as the hypothesis to prepare for the empirical part of this dissertation.

Chapter 8 constitutes the study that was conducted to investigate the research question and test the hypotheses that were formulated in the previous chapter. Initially, the methods that were applied are presented, a combination of behavioural data, event-related potentials, frequency band analysis as well as competence data. To substantiate the use of event-related potentials, their fundamentals are explained which is followed by a detailed description of relevant ERP components and a small literature review that validates the tasks that are used to examine the three main executive functions in the study as well as to give an overview of relevant areas of interest and electrodes that are used for the ERPs. Subsequently, the tasks that are used to test the executive functions are explained in detail, followed by an in depth description of the participants that took part in the study, the stimuli that were used, how the data was collected as well as how the data was analysed. Thereafter, the results of the study are presented, starting out with the competence data and followed by behavioural data as well as electrophysiological data.

The findings that were presented in Chapter 8 are critically discussed in detail in Chapter 9 where they are also contextualized within the framework of the relevant literature that that was presented beforehand and compared to previous studies. Chapter 10 marks the conclusion of the study, where all findings are again briefly summarized and put into context. Furthermore, the implications and relevance of the findings are discussed. Lastly, Chapter 10 also presents the limitations of the study as well as future research endeavours that are relevant to the research field in the context of this dissertation.

2. Neuroplasticity and Learning

There is little doubt that bilingualism and therefore managing multiple languages in one brain changes the brain both functionally and structurally, a hypothesis for which studies comparing monolingual and multilingual brains have provided ample evidence (see e.g. Mechelli et al. (2004); Felton et al. (2017); Burgaleta et al. (2016); Luk et al. (2011)). Yet, there seems to be something even more distinct about simultaneous interpreting (SI), a specific modality of bilingualism, that distinguishes interpreting brains from non-interpreting bilingual brains. Studies comparing interpreters and bilinguals or multilinguals have shown increased grey matter volume in different brain areas (e.g. Becker et al. (2016); Babcock (2015)), increased functional connectivity (e.g. Becker et al. (2016)) or increased subcortical structure volume (e.g. Korenar et al. (2023)) for interpreters compared to bilinguals. Furthermore, increased cortical thickness in regions related to speech processing and executive functions, thereby expanding into the non-linguistic realm (e.g. Hervais-Adelman et al. (2017)) have been shown for interpreters compared to multilinguals and sometimes translators. Interpreters have also been shown to outperform bilinguals in different executive functions tasks (Morales et al. (2015); Yudes et al. (2011); Henrard & van Daele (2017); Dong & Zhong (2017)). Therefore, the question of what it is about simultaneous interpreting that leads to adaptions in interpreting brains that are not apparent in bilingual brains is a relevant one.

Simultaneous interpreting requires bilingual speech processing, similar to written translation and speaking two languages in general, however, interpreting and especially simultaneous interpreting requires the near simultaneous processing and coordination of two languages at the same time (Ahrens 2011: 105–106). Highlighting this simultaneity and the overlapping processing during SI, Babcock & Vallesi (2017) explain that SI "[…] requires an individual to simultaneously comprehend speech in one language, transform the meaning into another language, and produce the resulting output" (Babcock & Vallesi 2017: 403). But first and foremost, simultaneous interpreting is a skill that is learned and requires the acquisition of explicit competences (see Section 6.2). This is also supported by Babcock & Vallesi (2017), who explain that SI is also a form of bilingualism and a skill that needs to be acquired

(Babcock & Vallesi 2017: 403). At the core of every acquisition, learning, training, and skill process lies neuroplasticity. It can be referred to as the brain's capacity to "modify itself, functionally and structurally, in response to experience and injury" (Bernhardi et al. 2017: 1). Learning how to manage multiple languages and specifically learning to interpret simultaneously can be seen as such an experience as mentioned by Bernhardi et al. (2017). However, before specifically focussing on how learning to manage multiple languages and finally how learning simultaneous interpreting can affect brain plasticity in executive functions, this chapter initially explains from a language-independent point of view the basis of how both structural as well as functional changes can arise as a function of learning and plasticity.

Generally speaking, there are two main kinds of plasticity. Functional plasticity involves changes in the functionality of neurons, such as the firing rate or the amount of neurotransmitters that are released. Structural plasticity entails changes in brain volume effected the formation of new neural connections (Costandi 2016: 13).
Already in 1948, the Polish neuro-physicist Jerzi Konorski deduced that the repetition of stimuli, by which he meant learning, can cause longer lasting transformations in certain neuron systems, which he called plastic changes, or, plasticity (Konorski 1948: 79–82). Further developing from this, Donald Hebb coined the Hebbian learning law in his book *The Organization of Behavior: A Neurophysiological Theory* in 1949. He assumed that:

> *"When an axon of cell A is near enough to excite a cell B repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased" (Hebb 1949: 62).*

While Hebb initially intended to define the cellular basis of learning in memory with this postulate, it formed the basis for long-term changes in synaptic strength. In other words, Hebb hypothesized that the simultaneous activity of presynaptic terminals and postsynaptic neurons reinforces their mutual synaptic connection. This suggests that recurring mutual activity not only strengthens the connections between neurons but makes it more likely that the connection is retained or that even new connections are formed. Correspondingly, synaptic terminals that do not receive as much activity are weakened and will in time lose their connection to the postsynaptic cell. This mechanism explains how the brain continues to grow after birth in reaction to the new experiences a newborn is subjected to. This decline of synaptic connection can especially be seen in adolescence, when only the most frequently used connections, namely those shaped by experience persevere (Purves et al. 2012: 537–538). This concept of strengthening connections by frequent use has been applied in translation studies in the context of the gravitational pull hypothesis by Halverson (2003). She drew on the concept of entrenchment coined by Langacker (1987) who determined that any cognitive occurrence or cognitive event leaves a neurochemical trace behind that is supposed to facilitate the reoccurrence of said event. He states that "[…] an event becomes more and more deeply entrenched through continued repetition" (Langacker 1987: 100). He furthermore explains that a deeply entrenched event can be carried out more automatically than events that have not been repeated often enough to become entrenched (Langacker 1987: 100). Halverson (2003) extrapolated this concept to the field of translation and assumes that in an activated linguistic network, some linguistic nodes are more prominent than other, among other things because of their deeper entrenchment caused by frequent activation (Halverson 2003: 216). She proposes that when translation equivalent routes have been repeated or activated several times, their connections are strengthened and therefore entrenched (Halverson 2003: 215). When translation routines are repeated often, they become more entrenched and automated (Halverson 2003: 200). In addition to entrenchment, salience also plays a role, a term Halverson (2003) originally referred to as "prominence" (Halverson 2003: 206–207), since some linguistic features are more prominent or salient than others. Halverson (2017) states that activation patterns that have been used more frequently over time and have become more entrenched are more prominent or salient than others, which causes the linguistic forms to be more likely to be chosen (Halverson 2017: 13). Halverson has described this as a "gravitational pull" (Halverson 2003: 216) and has also referred to this as "magnetism" (Halverson 2017: 14) in more recent publications since it better captures the concept that during translation, the translator is more drawn to target language items that exert more salience than others (Halverson 2017: 14)

It has become clear from this chapter that coactivation and the strengthening of connections, both from a general neuroanatomical view as well as from a linguistic view is the key for plasticity and ultimately for learning to occur. Importantly, strengthening synaptic connections and synaptic pruning can occur in all stages of life. They are a crucial part of the brain reshaping itself as a consequence of learning, which is explained further in the next subsection.

2.1 Learning-induced plasticity

In the beginning of brain research, it was assumed that the brain was incapable of changing. However, works such as those by Diamond et al. (1964) lead to a paradigm change that showed that the brain is capable of changing in reaction to external experiences. This change is initially marked by an increase in grey or white matter volume. Diamond et al. (1964) used rats to study the influence of experience and environmental enrichment on the brain. Postpartum, rats were put in one of two groups. One group received environmental complexity training, the other group was put in an isolated condition and was depraved of new experiences. Results showed an increase of depth of cortex, especially in the visual cortex bilaterally for the environmental enrichment group compared to the sensory deprivation group. Furthermore, the environmental enrichment group showed a decreased number of neurons, glia, and capillaries per field. This was interpreted as an indication of more intercellular and intervascular substance due to more dendritic branching (Diamond et al. 1964: 111–118). Later, studies in humans have also shown that changes in brain volume and thickness can occur in adults, particularly in brain regions that are responsible in function for the task that is executed or acquired (Lövdén et al. 2013: 2303–2304). Evidence for this comes for example from studies with taxi drivers or jugglers. Maguire et al. (2000) used MRI to compare the brains of London taxi drivers who have extensive navigation experience to those of a control group who were not taxi drivers. Results showed significantly increased grey matter volume in the posterior hippocampi for the taxi drivers compared to controls. This matches with the findings that the posterior hippocampus is implicated in recalling already learned spatial representations of the surroundings, especially in storing and using mental maps. Since the taxi drivers highly depend on their navigational skills, their hippocampus seems to have adapted to the increased demands (Maguire et al. 2000: 4398–4402). It must be said however that it is not clear from this study if there were no other confounding variables next to the navigational skills that might have led to changes in the hippocampus. Similarly, Draganski et al. (2004) compared grey-matter volume of two groups in a pre-and post-test using MRI. One group had to learn juggling within a three month period in which they regularly practiced juggling. The other group did not receive juggling training. Both groups were scanned before the beginning of training and at the end of training. Finally, the jugging group was scanned an additional time three months after their training ended; they were no longer fluent in the juggling skill. Results showed that only the juggling group displayed increased grey-matter volume in the bilateral mid-temporal area and in the left posterior intraparietal sulcus after the second scan. After the third scan, this expansion decreased again (Draganski et al. 2004: 311). Taken together, results from both studies show that the adult brain is capable of changing and that change is not only present after birth and is not only the result of brain aging but also related to learning.

The observation that acquiring and using a new skill may occur alongside structural adaptions in brain regions that are related to the execution of the skill is often referred to as experience-dependent plasticity (Pliatsikas 2020: 459). Draganski et al. (2004) also referred to the change of brain structure in adults in reaction to environmental demands as learning-induced plasticity (Draganski et al. 2004: 311). Therefore, based on Hebb's postulate that neurons that are active together form stronger connections, learning and learning to adapt to new environmental influences or experiences is seen as one of the biggest influences of plasticity in the brain.

However, what exactly constitutes the increases in brain volume is not entirely clear. It is assumed that it is a combination of changes in dendritic branches, synapses, cell numbers, cell sizes and capillaries (Lövdén et al. 2013: 2300).

Changes in volume and thickness of grey and white matter are assumed to reflect learning and skill acquisition, however, the time-course of volumetric changes in the brain is an important aspect to consider, since an increase in volume might not last. In turn, this does not have to imply that the learning gains are lost. Interesting evidence for this comes from Quallo et al. (2009), who used MRI to investigate the learning process of adult macaque monkeys who learned to use a rake. The monkeys were scanned before, during and after their intensive training. Results showed a learning related volumetric increase in several brain areas that corresponded to the learning curves of the monkeys. Monkeys that learned fast already showed a grey-matter increase after one week of training, while a monkey that showed a flatter learning curve showed later and slower grey-matter volume increased. A crucial point here is that when the monkey's performance reached a plateau, grey-matter volume decrease although the training continued. After training had finished, grey-matter volume was greater than at the beginning of training, but smaller than during the monkey's performance peak. The authors concluded that the increases in grey matter volume were due to the tool learning experience (Quallo et al. 2009: 18397-18383). Lövdén et al. (2013) called this phenomenon "the expansion-partial renormalization hypothesis" and assumed that a rapid expansion of brain volume in the initial phase of training succeeded by a renormalization of brain volume since only the most used new connections are maintained, might be a general principle of neural plasticity (Lövdén et al. 2013: 2301). This increase in grey matter volume followed by a decrease and renormalization phase has also been discovered in the field of language learning which is discussed in more detail as a theoretical basis for plasticity in simultaneous interpreters in 5.4.

Lövdén et al. (2013) emphasize that one should be cautious in equating an increase in brain volume with increased performance of the individual and that further research is necessary to establish for certain that learning is a prerequisite for volumetric brain changes or whether increased activity in relevant brain regions induced by the new experience suffice to show an increase in volume without any learning or improvement in performance (Lövdén et al. 2013: 2303–2304).

Apart from the expansion-partial renormalization hypothesis, skill learning has also been discussed on a more theoretical level, however on the same basis, namely Hebb's (1949) learning postulate (see Chapter 2.). Based on observations from brain research that have shown that over the course of learning, experience-dependent changes happen in the brain, Chein & Schneider (2012) explore the time line of changes in the brain during learning. Grounded in the Hebbian learning postulate, Chein & Schneider (2012) present a triarchic theory of learning and skill acquisition. They explain how specific brain networks change as an effect of different stages of learning. Within a hierarchically organized learning system consisting of a (1) representation system, the (2) cognitive control network and the (3) metacognitive system, the learner lives through three stages of learning, beginning with the formation or acquisition of new behaviour that is converted into a controlled execution stage that ultimately results in a stage of automatic execution. While learners acquire and execute new skills, activity in these three systems change according to the learning stage. While activity in the cognitive control network is low in the formation stage, the activity of controlling and executing the new skill increases in the controlled execution stage of learning (Chein & Schneider 2012: 78–80). This three-phase model of learning and skill acquisition becomes more important later in this dissertation in the context of interpreting competence acquisition (see Section 6.2) as well as in the context of structural and functional brain changes as an effect of managing multiple languages (see Chapter 3).

Chein & Schneider (2012) explain that the representational system acts as an acquisition and storage system for memories and basic knowledge and creates associations based on people's experiences, serving as an archive for accumulated knowledge. The learning mechanism of this system is said to rely on the Hebbian postulate that cells that often fire simultaneously create stronger connections (see Chapter 2.). Eventually, the representational system can process information without the cognitive control network which leads to faster and ultimately automatic responses to stimuli. This is why its activity is low in the first learning stage of formation, augments in the controlled execution stage and reaches a high in the automatic execution stage. Within their model, Chein & Schneider (2012) describe the cognitive control network as a "domain-general system responsible for monitoring and directing information processing in the representation system" (Chein & Schneider 2012: 81). This is the first mention of a domain-general executive control system in this context of skill acquisition and learning and will become more important in Chapter 4, when executive functions are explained, in Section 5.3 when the overlap of executive functions and language processing is laid out and in Section

6.2, when the acquisition of interpreting competences are examined. Within the hierarchic model, the cognitive control network is responsible for focussing attention on the relevant information and collaborates with working memory so that people can uphold their goals and keep out the irrelevant information. Its most important node is the lateral prefrontal cortex. In contrast to the representation system that only stores explicitly learned associations, the cognitive control network represents more complex information. The cognitive control network is highly involved when new information is processed in the representation system by controlling the focus of attention so that it focusses on the important information and the task goal can be maintained. Over time, the representation system becomes familiar with thoroughly studied associations and can tap directly into the information they contain. Therefore, the cognitive control system no longer has to give permission and is less activated, the representation system can deal with the information independently during the stage of automatic processing. Chein & Schneider (2012) mention that training can play a significant role in modelling and strengthening cognitive control capacity, as the regions of the cognitive control network become increasingly efficient. Furthermore, Chein & Schneider (2012) summarize that different cognitive control processes can be enhanced by different types of experiences. This means that training cognitive control can enhance, for example, the ability to update active mental representations, to multitask or to suppress interferences from distractions (Chein & Schneider 2012: 81–82).

Interestingly, these descriptions of control processes that can be enhanced corresponds to the main domain-general executive functions discussed in detail in Sections 4.1, 4.2 and 4.3, inhibition/interference suppression, monitoring, working memory updating and shifting. The involvement of interference suppression, working memory updating and shifting in bilingual language processing in general has been discussed extensively, as laid out in further detail in Sections 5.1 and 5.2. Furthermore, the involvement of these processes in translation and especially simultaneous interpreting has a special status. For example, monitoring and inhibition have both been assumed to be involved in conflict monitoring for interference of a more activated language or monitoring of speech input and output during simultaneous interpreting (for details see Chapter 6).

Additionally, the metacognitive system can be regarded as a control system that oversees the cognitive control network. It is involved in establishing new behavioural routines, prepares and initiates tasks and monitors them (Chein & Schneider 2012: 82). The metacognitive system plays its most important role in the formation stage, where it is responsible for initiating and preparing the brain to carry out a novel routine. This is followed by the controlled execution stage, where the cognitive control network guides the attention to the most important information and prevents distractions. Ultimately, when the system has learned to automatically execute the information, the representation system is most active, the cognitive control system can let go to some degree since the association between the stimulus and the appropriate response has been sufficiently established (Chein & Schneider 2012: 83).

Lastly, Chein & Schneider (2012) assign the different phases and systems to brain regions. They associate the formation stage with an increase in activity in the anterior prefrontal cortex that they ascribe to the metacognitive system. When the skill acquisition shifts into the controlled execution stage that revolves around the cognitive control network, Chein & Schneider (2012) describe that the activity in the anterior prefrontal cortex diminishes and in turn grows in the dorsolateral prefrontal cortex, the anterior cingulate cortex, the posterior parietal cortex and the inferior frontal junction. When the skill is mostly automated after practice, the activity in the cognitive control areas is reduced, task performance is mainly linked to the representation system (Chein & Schneider 2012: 81). Note that the cognitive control areas mentioned by Chein & Schneider (2012) largely correspond to the executive control network established by Abutalebi & Green (2007; 2008) and Green & Abutalebi (2013) in Section 5.3.

This subchapter has explained the concept of experience- or learning dependent plasticity in the adult human brain by presenting studies on animals and humans that rely on the Hebbian learning postulate laid out in Chapter 2. Furthermore, two important frameworks have been discussed that examine how learning a skill affects the time course of experience-dependent plasticity on a structural level (expansion-partial renormalization hypothesis) and that explains how a new skill is learned on a more theoretical level (hierarchical model of learning). The latter framework explicitly incorporates the role of domain-general executive functions, which provides an important basis for the following chapters that explain how managing multiple languages and especially simultaneous interpreting can impact the brain and executive functions.

In preparation of domain-generality, the next subsections briefly highlight another aspect of plasticity, namely the transfer of a skill learned in one domain to another domain that was not specifically trained.

2.2 Transfer of expertise

As laid out in the previous subsection, learning and skill acquisition can impact the structure and functionality of our brain. Interestingly, this reorganization is not necessarily limited to the domain that was trained, but can be transferred. Kimball & Holyoak (2000) see transfer as a basic process of learning and explain transfer as the performance of a task that affects the performance of a second task. They argue that transfer and expertise are related in such a way that certain types of expertise allow for a transfer of learning to new tasks. However, in order for transfer to occur, two tasks must share features that are relevant for achieving the goal of the tasks. According to their reasoning, experts can better transfer knowledge to a new task that even lies beyond their initial domain of expertise (Kimball & Holyoak 2000: 109–118).

Although they do not specify the features that must be shared between two tasks so that a transfer of expertise can occur, it is possible that the common denominator here are domain-general executive functions (see Chapter 4). Evidence for this comes from various studies that show a successful skill transfer in the field of musical expertise as well as video games: In a study with children aged 6-15, Ho et al. (2003) investigated how music training affects visual and verbal memory. One group was in a musical programme at school and received additional training. The other group had no such training and had not participated in a musical programme before. Both groups were matched on intelligence and subjected to different tasks of verbal and visual memory. As a result, children with music training showed better verbal memory than the group without music training. Schellenberg (2004) conducted a study with four groups of children: Participants in two of the groups received instrumental or voice training, participants in two other groups either received no training at all or were trained in drama, a non-musical artistic activity. The IQ was measured before and after the training period. Children in the music groups showed greater increases in IQ compared to the control groups. Interestingly, children in the non-musical artistic activity groups showed improvements in adaptive social behaviour. In an ERP-study, Moreno et al. (2011) trained two groups of children in musical curriculum or a visual arts curriculum. An intelligence test as well as an executive functions task were administered as pre- and post-tests. The results showed enhanced performance in verbal memory as well as an improvement in executive functioning. Results also revealed an increased P200 ERP amplitude that has been interpreted to reflect an increased neural representation of the executive function skill by training. P200 refers to a positive peak of electric potential 200ms after a stimulus has been presented.

Similarly, studies have investigated how performance in video games can be transferred into the domain of cognitive functions. Buelow et al. (2015) observed the effect of short term video game playing and compared it to a control group. Participants in the video game group showed improved performance on measures of decision making and problem solving, and made fewer errors, reflected in the Wisconsin Card Sorting test. In four experiments, among them a Flanker compatibility task, Green & Bavelier (2003) showed changes in various aspects of visual attention when they compared frequent video game player with non-players.

Taken together these studies and the knowledge on brain plasticity, learning and transfer of skills show that our brain and especially our cognitive control functions are exceedingly receptive to different experiences. One of these experiences that can also massively impact our cognition is bilingualism, controlling more than one language. The following chapter will therefore initially try to gauge the multifaceted term of bilingualism and then provide evidence on how managing multiple language can induce structural and functional plasticity in the brain. Studies from both bilingualism research as well as translation studies research will be provided.

This rather general approach to plasticity, skill acquisition and transfer of expertise however also translates to the linguistic realm. The following subsections will present experimental evidence from studies that have shown that learning to manage multiple languages in one brain is one of the external experiences that show structural as well as functional effects on both linguistic as well as non-linguistic executive functions. Transfer of expertise and training domain-general executive functions is also not a phenomenon that is restricted to the bilingual research domain. Learning to translate and to interpret between multiple languages and acquiring interpreting and translation specific competences also requires skill acquisition that is assumed to extend to the executive functions domain as will be explained in more detail in Chapter 6.

3. The plastic bilingual brain

The term bilingualism, although conveying the impression of being clear-cut, is much more complex and requires careful differentiation. More limited definitions of the term have described bilingualism for example as "[…]native-like control of two languages" (Bloomfield 1933 // 1984: 56), using only the proficiency level of speakers as indicators of bilingualism. However, this definition neglects important aspects of the experience of managing two languages and leaves open the question of which competence level would exactly classify a speaker as bilingual. In a more recent and broader definition, Grosjean & Li (2013) describe bilingualism as "[…] the use of two or more languages (or dialects) in everyday life" (Grosjean & Li 2013: 5). This however leaves room for speculation about the frequency, context, and modality of use of two languages in everyday life. When discussing bilingualism in depth and especially when conducting research with bilinguals and monolinguals, it should be classified in accordance with various dimensions, such as age of acquisition (AoA), learning context, sociocultural and linguistic environment, context of language use, detailed differentiation of language competence according to dominance, fluency and literacy of language, speaking and finally switching habits. AoA is often grouped into the subcategories early simultaneous bilinguals for a parallel acquisition of both languages before the age of three, early sequential bilinuals for instances in which the acquisition of the second language before the age of six and late sequential bilinguals for instances where the second language is usually learned after the age of three or six. Language competence is often grouped into balanced and unbalanced bilinguals. The first case covers bilinguals that have obtained the same level of proficiency in both languages while the latter case covers bilinguals that show dominance in one language (Schwieter & Festman 2023: 6–7). However, classifying bilingualism is not a trivial endeavour and still poses difficulties for bilingualism researchers everywhere. As Marian & Hayakawa (2021) explain, "[…] attempting to quantify 'how bilingual' is an individual (especially relative to others) is particularly challenging due to the many forms that bilingualism can take" (Marian & Hayakawa 2021: 527). Although there are standardized tests and self-reports such as the Language and Proficiency Questionnaire (LEAP-Q) (Marian et al. 2007) that attempts to generate a reliable measure of bilingualism or the Bilingual Switching Questionnaire (Rodriguez-Fornells et al. 2011) that tries to capture the extent to which bilinguals switch between languages in different contexts, there are still caveats. Marian & Hayakawa (2021) summarize that especially self-reported measures of proficiency can vary across skills, measures, ages, proficiency levels and raters and explain the option of combining self-reports with standardized tests of the bilingual ability such as picture naming or different vocabulary tests. They highlight that a prominent issue in bilingualism research is that the methods and measures that are applied to characterize bilingual and monolingual participants are still extremely heterogenous and that participants are often still grouped roughly together into a bilingual and a monolingual group, although there is a common consensus that bilingualism should be seen as a continuum. Most studies presented in this dissertation share this issue that is inherent to the whole research field. Populations often differ greatly in factors such as language history and language experience which, in addition to the lack of standardization in methods, makes it difficult to interpret results across studies. Marian & Hayakawa (2021) emphasize that "Until there is greater consensus on how individual differences should be assessed, it will be challenging to determine whether inconsistent results across studies emerge from differences in populations or measures" (Marian & Hayakawa 2021: 534). This is especially important for studies researching the bilingual advantage (see Section 5.4). Marian & Hayakawa (2021) recommend that researchers collect detailed participant information and try to establish whether there are relationships across some variables, as well as using mixed-effect models that can account for variance due to participant or item influences (Marian & Hayakawa 2021: 527–534). They even propose computation of a bilingual quotient, similar to the intelligence quotient but acknowledge the issues that come with the depiction of intelligence in the IQ (Marian & Hayakawa 2021: 542). Translation and Interpreting research are challenged by similar issues, as explained in more detail in Section 6.1. The measure of translation or interpreting competence is extremely important in experimental research, since participants also need to be classified in terms of their abilities in applying certain translation or interpreting skills. Bilingual competence is a part of translation and interpreting competence (see e.g. PACTE group et al. (2003), EMT framework European Commission 2022)) which makes it all the more important to establish standardized and comprehensive measures when studying translators and interpreters (see Sections 6.1 and 8.1.3).

Coming back to the bilingual brain, it is most important in the context of cognitive studies that the bilingual brain is seen as a complex and plastic entity that is formed depending on various language modes and experiences. As noted early on by Grosjean (1989), bilinguals and therefore the bilingual brain should not be seen as "two monolinguals in one person " (Grosjean 1989: 5), but as "an integrated whole which cannot easily be decomposed into two separate parts" (Grosjean 1989: 6) with languages interacting in complex neural networks.

In the context of this dissertation, bilingualism is seen as the act of managing more than one language in the brain. More specifically, the focus lies on the cognitive mechanisms for using and controlling two languages and how these mechanisms are affected and altered by simultaneous interpreting training, a form of translation activity and an extreme case of language control within the bilingual experience. The following subsections of Chapter 3 consider neurolinguistic studies on how bilingualism changes the structure and functionality of the brain. Section 3.1 presents studies that approach this topic from the viewpoint of bilingualism studies while Chapter 3.2 presents studies that approach this topic by from a more translation and interpreting-oriented perspective.

3.1 Language-induced brain plasticity in bilingualism research

The action of managing more than one language in the brain has long been assumed to affect the architecture and functionality of our cognitive abilities. Managing two language was initially thought to be non-beneficial, it was even presumed to negatively impact language development in children. However, a study by Peal & Lambert (1962) was the first to show a positive impact of bilingualism: bilingual children performed better than monolingual children on tests that required symbol manipulation and reorganization. In terms of increased executive functions in one French-English bilingual participant they observed that "Intellectually his experience with two language systems seems to have left him with a mental flexibility, a superiority in concept formation, and a more diversified set of mental abilities" (Peal & Lambert 1962: 20).

Since then, the field of neurolinguistics has come a long way in examining the impact of acquiring and using more than one language, or more specifically, how the brain restructures in order to accommodate linguistic representations of more than one language and control the processing efforts this entails.
As mentioned in Chapter 2, acquiring a new skill restructures the brain which can manifest in higher grey or white matter brain volume as well as increased functional connectivity. In the context of language acquisition and bilingualism, the main findings concern both cortical grey matter and subcortical structures, the cerebellum, white matter structural connectivity as well as resting state functional connectivity. Mechelli et al. (2004) compared monolinguals to early and late bilinguals, revealing higher grey-matter density in the left posterior supramarginal gyrus (pSMG) that was more pronounced for early than for late bilinguals. The inferior parietal cortex is involved in verbal fluency tasks (Mechelli et al. 2004: 757). Grogan et al. (2012) examined the grey matter volume of bilinguals and multilinguals and found increased grey matter density in the pSMG for the multilingual group (Grogan et al. 2012: 1350). Olulade et al. (2016) compared grey matter volume of monolinguals to unimodal and bimodal bilinguals and found increased volumes in the left hemisphere in the dorsolateral prefrontal cortex (DLPFC) with the middle frontal gyrus (MFG) and the precentral gyrus (PCG), the inferior frontal gyrus (IFG), inferior middle and superior occipital gyri, cuneus as well as posterior middle temporal gyrus (MTG) for unimodal bilinguals. In the right hemisphere, the IFG, MFG, superior frontal gyrus (SFG), frontal operculum, inferior parietal lobule (IPL), MTG and superior temporal gyrus (STG) showed increased volume for unimodal bilinguals (Olulade et al. 2016: 3199). Felton et al. (2017) also compared grey matter volume of monolinguals and bilinguals and found greater volume of the right anterior cingulate gyrus and sulcus, the pars triangularis, the inferior frontal sulcus as well as the superior temporal gyrus for bilinguals. Interestingly, they also observed a volume increase in corpus callosa in the central callosal region (Felton et al. 2017: 4–6). These studies show that the experience of bilingualism and controlling more than one language induces cortical

changes in areas that manage phonological and lexico-semantic processing but also language and domain-general executive control (Pliatsikas 2019: 235). Especially the anterior cingulate cortex hosting the anterior cingulate gyrus has often been correlated to domain-general executive control (Abutalebi et al. 2012).

Subcortically, Hosoda et al. (2013) found increased volume of the caudate nuclei in a vocabulary learning study after a learning period of 16 weeks (Hosoda et al. 2013: 13666). Burgaleta et al. (2016) compared monolinguals to simultaneous bilinguals and found greater volume of the putamen, thalamus, globus pallidus as well as right caudate in bilinguals (Burgaleta et al. 2016: 440). Similar results were provided by Pliatsikas et al. (2017), who examined bilinguals who differed in immersion and L2 usage with monolinguals. Results showed bilateral increased volume in the putamen, globus pallidus and right thalamus. Furthermore, immersion was a predictor of increased volume in right globus pallidus (Pliatsikas et al. 2017: 1788).

Concerning white matter changes, the most interesting structure in the context of this dissertation is the corpus callosum (CC), which ensures among others, interhemispheric communication with regions of the prefrontal cortex such as inferior frontal regions (Lacoste et al. 1985: 582). The prefrontal cortex has also been shown to be highly involved in both language control and executive functioning (Abutalebi & Green 2008, 2016). Coggins et al. (2004) compared monolingual teachers to bilingual teachers and found increased volume in the anterior midbody of the corpus callosum in bilinguals (Coggins et al. 2004: 72). Luk et al. (2011) compared monolinguals to lifelong bilinguals and found increased white matter density bilaterally in the superior longitudinal fasciculus (SLF), right inferior fronto-occipital fasciculus (IFOF) and the right uncinate fasciculus (Luk et al. 2011: 16810). Nichols & Joanisse (2016) correlated L2 proficiency and AoA in bilinguals with white matter density changes and found positive correlations between AoA of L2 and increased density in the right inferior longitudinal fasciculus (ILF), arcuate fasciculus and the corpus callosum, as well as correlations between L2 proficiency and increased density on the right ILF, arcuate fasciculus and forceps minor that belong to the corpus callosum (Nichols & Joanisse 2016: 20). Rossi et al. (2017) investigated the impact of L2 proficiency levels on white matter changes and compared monolinguals to proficient but late L2 learners. They found higher white matter density in the corpus callosum, IFOF and inferior longitudinal fasciculus (Rossi et al. 2017: 5–6). Taken together, these studies show that managing more than one language affects various white matter areas that are involved in linguistic processing, for example in semantic, syntactic, and phonological processing on the one hand, and areas implicated in executive functions on the other hand. Since managing two languages is thought to pose increased demands on language as well as cognitive control, myelination of white matter tracts is elevated, resulting in higher white matter density to ensure effective communication between both hemispheres (Pliatsikas 2019: 242).

As for the impact of managing multiple languages on brain functionality, most studies focus on changes in resting state connectivity, which is the connectivity of brain regions when no task is to be performed. Berken et al. (2016) used fMRI to compare resting state connectivity in sequential to simultaneous bilinguals with focus on the inferior frontal gyrus. They found stronger functional connectivity for simultaneous bilinguals between the left and right inferior frontal gyrus and between the inferior frontal gyrus and the dorsolateral prefrontal cortex, inferior parietal lobe and cerebellum, all regions partly involved in language and/or executive control. Especially the right inferior frontal gyrus plays a role in domain-general executive functions such as response inhibition and attentional control. Furthermore, L2 AoA was correlated with strength of functional connectivity, with earlier L2 AoA correlated with stronger functional connectivity. They also found that during speech production, the left inferior frontal gyrus showed reduced activation despite increased functional connectivity. Taken together, the results were interpreted towards increased neural efficiency for simultaneous bilinguals and for bilinguals with earlier L2 AoA.

Similarly, Kousaie et al. (2017) also compared resting state functional connectivity in simultaneous bilinguals to sequential bilinguals with focus on the ventromedial prefrontal cortex (VMPFC) and additionally tested the participants on a Simon task to assess the executive function of interference suppression. Results revealed that the VMPFC activity was correlated positively with areas of the so called default mode network, which show a decrease in activity when a task is present. So, as activity in the VMPFC decreased, activity in the default network also decreased which is associated with better performance of executive function tasks. Furthermore, the authors also found a negative correlation of activation in the VMPFC with areas of the task-positive attention network that show an increase in activation when a task is present. This means as activity in the VMPFC or the default mode network decreases, activity in the task-positive attention network increases, which is involved in important executive functions such as attentional control. Ultimately, results revealed that this anticorrelation between the default mode network and the task-positive attention network is influenced by the degree of bilingualism, with stronger anticorrelations between the VMPFC and the task-positive attention network for simultaneous bilinguals than for sequential bilinguals in combination with better interference suppression performance in the Simon task for simultaneous bilinguals.

Instead of applying a binary approach to bilingualism, Sulpizio et al. (2020) investigated the effect of L2 AoA, L2 proficiency and usage as continuous variables on resting-state functional connectivity in regions identified to be involved in language processing and executive functions, namely the prefrontal cortex, bilateral ACC, inferior parietal lobules, thalamus, basal ganglia and cerebellum (Abutalebi & Green 2008, 2016). They found that higher AoA values were associated with more functional connectivity between the left pSTG and the left praecuneus. Furthermore, an increase in L2 proficiency was associated with an increase in connectivity only for late bilinguals. Specifically, for regions associated with control, increased L2 proficiency was associated with increased efficiency. Furthermore, living in a bilingual context where both languages are used equally was associated with enhanced connectivity.

3.2 Language-induced brain plasticity in cognitive translation and interpreting research

In contrast to bilingualism research, cognitive translation and interpreting studies "[…] focus on the ways people translate and interpret oral, written and signed languages, in order to improve production, quality, professional training and user education" (Muñoz Martín 2022: 1). Evidence from the impact of managing more than one language in the context of translating and interpreting mostly stems from studies on interpreters and professional translators who were compared to bilingual or multilingual controls. A more in depth review of the research landscape of changes to the brains of interpreters specifically is provided in Section 6.3. However, the following studies represent selected contributions from studies that are more oriented towards translation and interpreting than towards bilingualism.

Elmer et al. (2010) examined auditory word processing in simultaneous interpreters with more or less interpreting experience and compared them to bilingual controls. Within the scope of an EEG study, participants performed a lexical decision task. Results showed larger N400 in interpreters during incongruent trial detection within the L1 and L2, as well as when they performed the task in the other direction from L1 to L2. These results were interpreted as sensitivity to semantic processing with and across L1 and L2 that was induced by interpreting training (Elmer et al. 2010: 152).

Elmer et al. (2011) compared white matter density of simultaneous interpreters to non-interpreting bilingual controls and found reduced density in various parts of the corpus callosum for the simultaneous interpreting group. They interpreted the results as expertise-related modifications of white matter that are due to special demands of simultaneous interpreting (Elmer et al. 2011: 2069). Elmer et al. (2014) compared grey matter volumes of professional simultaneous interpreters to multilingual controls. In contrast to the results in bilingualism research, simultaneous interpreters showed reduced grey matter volume in the left middle-anterior cingulate gyrus, bilateral middle-anterior insula, left SMG, bilateral pars triangularis and left pars opercularis compared to the multilingual control group (Elmer et al. 2014: 183).

Contrary to these findings, Becker et al. (2016) compared simultaneous interpreters to professional translators and found increased grey matter volume in the left frontal pole region for simultaneous interpreters, a region belonging to the prefrontal cortex associated with multitasking. Furthermore, they found increased functional connectivity between the frontal pole and the left IFG, pars opercularis and pars triangularis, as well as the left MTG in simultaneous interpreters. They concluded that interpreters show a network that is functionally more strongly connected than in controls (Becker et al. 2016: 257).

As a part of her dissertation, Babcock (2015) examined changes in grey- and white-matter in simultaneous interpreting students and translation students at the beginning and at the end of their two-year Master’s programme. She found an increase in grey-matter volume for interpreters in regions that are implicated in simultaneous language input and output. Furthermore, results showed that white matter language pathways in the left hemisphere were reinforced (Babcock 2015: 110).

Hervais-Adelman et al. (2017) measured cortical thickness in interpreting students before and at the end of their Master’s programme and compared both instances to multilingual controls. They found increases in cortical thickness in regions involved in speech processing and executive functions for the simultaneous interpreters compared to the controls: in the left posterior STG, SMG, the planum temporale within the Sylvian fissure and in the right hemisphere in the superior parietal lobule, the angular gyrus of the IPC, in the intraparietal sulcus and in the superior frontal gyrus (Hervais-Adelman et al. 2017: 215).

Korenar et al. (2023) compared volumes of the caudate nucleus and the putamen of translators, interpreters and bilinguals. They found that translators and interpreters showed significantly larger putamen and caudate volumes than bilinguals. However, there were no significant differences in putamen and caudate volumes between interpreters and translators. Interestingly, the authors also calculated the effects of bilingual experiences as indexed by a score on caudate and putamen volumes. For interpreters, volumes of both caudate nuclei initially increased with growing bilingual experience only to decrease after reaching a certain level of bilingual experience. For translators, only the volume of the caudate nucleus showed an initial increase and a following decrease in volume with growing bilingual experience. They reasoned that different brain regions adapt as a reaction to differences in bilingual experiences (Korenar et al. 2023: 5).

Since these studies compared interpreting individuals to non-interpreting individuals (bilinguals or translators) and found differences in different parts of brain architecture, this seems to suggest that there is something “special” or unique about how the practice of interpreting shapes the brain that even transcends the control demands that are put on managing multiple languages in “only” bilingual brains. There even seems to be a difference between translating and interpreting individuals, although not all studies confirm this difference (see Korenar 2023). This reinforces the question of what it is about the modality of simultaneous interpreting that warrants these changes; this is discussed further in Chapter 6. Points that need to be critically reflected in studies from bilingualism as well as translation and interpreting research mostly concern the issue of measuring bilingualism or bilingual proficiency as well as interpreting or translation competence that was already discussed in the introduction of Chapter 3. Only more recent studies try to gauge the impact of variables such as AoA or L2 proficiency or immersion (see e.g. Pliatsikas et al. (2017) or Sulpizio et al. (2020)). In the studies involving interpreters or translators and bilinguals, non-interpreting bilinguals were not examined on some important variables, for example on how much they switch between language which could impact results. Furthermore, it is not straightforward to interpret the opposing results of increased grey-matter volume in one study (Becker et al. 2016) and decreased grey-matter volume in another study (Elmer et al. 2014). Since measurements were often not accompanied by tests of performance, it is difficult to infer from more or less brain volume to better or worse performance or a more or less efficient brain. After all, most studies presented here apart from one study by Hervais-Adelman et al. (2017) or Babcock (2015) were cross-sectional with results only representing one moment in time.

Despite the criticism, these neurolinguistic studies from both bilingualism research and cognitive translation and interpreting studies show the general capability of the brain to adapt in response to different characteristics of managing multiple languages such as the age at which the languages are required, the proficiency of both languages as well as the modality of language use (e.g. simultaneous interpreting vs. bilingual language use in everyday life). This chapter has already briefly touched upon the fact that, during language processing, both linguistic brain regions as well as brain regions responsible for domain-general executive functions are affected by structural and functional changes induced by managing multiple languages. The following Chapter 4 provides a thorough definition of executive functions and describes selected executive functions that are prevalent in literature as well as relevant to this dissertation. Chapter 5 will then consolidate executive functions and their role in language processing as a preparation for discussing the effects of simultaneous interpreting on executive functions.

4. Executive functions

The terms cognitive control and executive functions have often been used interchangeably in research. It can be argued that cognitive control refers to a more general spectrum of mental processes, as laid out in Botvinick et al. (2001):

> *"A remarkable feature of the human cognitive system is its ability to configure itself for the performance of specific tasks through appropriate adjustments in perceptual selection, response biasing, and the on-line maintenance of contextual information. The processes behind such adaptability, referred to collectively as cognitive control […]" (Botvinick et al. 2001: 624).*

Executive functions have been defined as "[…] a collection of top-down control processes used when going on automatic or relying on instinct or intuition would be ill-advised, insufficient, or impossible" (Diamond 2013: 136). More precisely, they can be viewed as "[..] the ability to coordinate thoughts and actions in relation with internal goals a variation of top-down mental processes that help us coordinate our thoughts and actions […]" (Koechlin et al. 2003: 1181). Executive functions play a vital role in our everyday life since they are involved in higher cognitive processes such as planning and reasoning. Exercising executive control is seen as effortful, since it is harder to change and interrupt a process than to automatically continue down a path (Diamond 2013: 136; Koechlin et al. 2003: 1181). Early models such as by Norman & Shallice (1986) have already distinguished between more automatic processing that is carried out by learned and familiar action sequences called schemas that can be triggered by an event, and the supervisory attentional system (SAS) that controls attention. As opposed to automatic processing, the SAS is required if a situation is new and no automatic task schemas can be applied, the task goal needs to be changed, something needs to be suppressed or the task is difficult (Norman & Shallice 1986: 6–8). Miller & Cohen (2001) described automatic behaviours such as reacting to an unexpected stimulus form over time as a learning mechanism and can be classified as bottom-up processing. Bottom-up processing behaviour can therefore be categorized as being determined by external stimuli. In contrast to that, they describe the requirement of controlling behaviour by inner goals as top-down processing. Top-down processing behaviour is therefore not guided by external stimuli but by internal intentions. They explain that the prefrontal cortex is especially involved in top-down processing, namely in situations where automatic processing is not possible since more control is required, e.g. when we need to form internal goals and want to achieve them or the goals change frequently (Miller & Cohen 2001: 168).

As one of the most complex functions that can be carried out by the human brain, understanding all components of executive functioning is still an ongoing investigation. However, different attempts to categorize executive function have been undertaken, such as by Miyake et al. (2000). They have described executive functions as "[…] general-purpose control mechanisms that modulate the operation of various cognitive subprocesses and thereby regulate the dynamics of human cognition" (Miyake et al. 2000: 50). They defined three "core" executive functions: inhibition, shifting (and monitoring) and working memory updating. In their "unity and diversity model", they show that although executive functions do have distinctive aspects and are separable from one another to a certain extent, they also share one important common factor that can be considered the core of executive functioning: the ability to maintain task goals (Miyake & Friedman 2012: 10–12).

Although that it is beyond question that executive functions involve neural networks that extend beyond frontal regions, the prefrontal cortex has been identified as particularly important for executive functioning (Miller & Cohen 2001: 168). Studies such as a PET study by Collette et al. (2005) provided neuropsychological evidence for the distributed network of executive functions. Adapted to the unity and diversity model by Miyake et al. (2000), they examined which brain areas are commonly active during all three main executive functions tasks and which brain area activation is task specific. They found that all three tasks shared the activation of the left superior parietal cortex, the right intraparietal sulcus, the left middle frontal gyrus as well as the left IFG. The right supramarginal gyrus, the left praecuneus, left superior parietal cortex as well as the left middle and inferior frontal gyrus was specifically associated with switching, while updating specific areas included the frontopolar, superior, middle, inferior and orbitofrontal cortex as well as the intraparietal sulcus and the cerebellum. Finally, inhibition mostly showed activation on the commonly activated cerebral areas, however, the right orbitofrontal gyrus and the right middle/superior frontal gyrus seemed to be more associated with inhibition than with switching or updating (Collette et al. 2005: 418–419).

A more detailed account on the location of executive functions can be found in Section 5.3.

Following this definition of the broad terminus executive functions, the next subchapters dive into the three main executive functions as classified by Miyake et al. (2000) in greater detail.

4.1 Inhibition

Miyake et al. (2000) describe inhibition as the ability to wilfully inhibit dominant and automatic responses (Miyake et al. 2000: 57). Later, this definition of inhibitory control was extended to include controlling attention, behaviour as well as thoughts and emotions in order to follow an internal goal (Diamond 2013: 137). Furthermore, researcher have proceeded to distinguish between two kinds of cognitive control that are often grouped together under the term inhibition: interference suppression and response inhibition (Bunge et al. 2002a; Martin-Rhee & Bialystok 2008; Esposito et al. 2013; Luk et al. 2010). Interference suppression is the ability to filter out irrelevant information surrounding a stimulus, whereas response inhibition is characterized by inhibiting responses that are inappropriate but dominant (Bunge et al. 2002a: 301). Bunge et al. (2002a) conducted an fMRI study that examined the brain activation in these two kinds of inhibition in children compared to adults by means of a combination of Flanker tasks and a go/no-go paradigm. The experiment contained three relevant conditions. In a congruent condition, participants saw five arrows in a row all pointing in the same direction and had to press a button to indicate the direction of the five arrows. In the incongruent condition, participants saw five arrown ina row with the central arrow pointing in a different direction than the others. Participants had to press a button to indicate the direction of the central arrow, a task requiring the inhibition of the interfereing information from the arrows surrounding the central arrow. In the no-go condition, participants saw X’s instead of arrows and had to inhibit their response alltogether. On the one hand, the results showed that children engaged different brain regions from adults during interference suppression, indicating a developing cognitive strategy from childhood to adulthood. On the other hand, children seemed to activate a subset of the adult's response inhibition circuit. Taken together, these results point towards two different forms of inhibitory control with two different developmental timelines.

The notion of inhibitory control has also been associated with attentional control since we can actively choose on what to focus and supress the urge to give attention to other stimuli. Instead of letting external stimuli such as loud noise draw our attention away (involuntary automatic attention), we purposefully focus our attention. This is comparable to a cocktail party, where we focus on the conversation we are having at the moment and block out all the surrounding conversations. In this case researchers speak of voluntary attention that is goal driven and endogenous (Diamond 2013: 137). Posner & DiGirolamo (1998) speak of executive attention. For them, executive attention is at the root of executive control, and they use the term synonymously to executive control.

Inhibitory control can be tested by means of various psycholinguistics and psychological tests such as the Simon task (Simon & Wolf 1963), the Flanker task (Eriksen & Eriksen 1974), the Stroop task (Dunbar & MacLeod 1984), the attention network task (ANT) (Fan et al. 2002), go/no-go paradigms and the antisaccade tasks (Hallett 1978). Specifically, interference suppression and conflict monitoring are often tested with the Eriksen Flanker or the Stroop task, whereas response inhibition is often tested with the go/no-go paradigms. This dissertation focusses on the Flanker task to investigate the changes in inhibition and conflict monitoring during interpreting training. The Flanker task is one of the classic stimulus-response-conflict (SRC) tasks often applied in cognitive psychology. In conflict tasks, interference surfaces because attributes that are not relevant to a stimulus are integrated in a stimulus that also contains relevant attributes (Kleinsorge 2021: 2). In cases of the Flanker tasks, the stimuli are classically letters or arrows in congruent and incongruent conditions. The Flanker task has widely been used to examine the bilingual advantage in executive functions in bilinguals (see Section 5.4) as well as in translators and interpreters (Section 6.4).

A concept that is closely related to inhibitory control or interference suppression is conflict monitoring. Botvinick et al. (2001) explain conflict monitoring as a process "[…] that monitors for the occurrence of conflicts in information processing […]" (Botvinick et al. 2001: 625). They explain that when conflict is detected during conflict monitoring, this information is passed on to centres that regulate executive control which result in compensatory actions such as interference suppression or inhibition. They also suggest that the anterior cingulate cortex (ACC) serves as the primary conflict detection area in the brain (Botvinick et al. 2001: 625). Monitoring, inhibition and interference suppression are further discussed in the context of bilingual and monolingual speech processing in Sections 5.1 and 5.2, in the context of the bilingual advantage hypothesis as well as code-switching that are discussed in Section 5.4 and furthermore in the context of cognitive interpreting processes discussed in Chapter 6.

4.2 Shifting/Switching

According to the definition by Miyake et al. (2000), the executive function shifting means shifting between different mental sets or tasks and is also referred to as task switching or attention switching. Most importantly, shifting requires disengaging an irrelevant task set and engaging a relevant task set (Miyake et al. 2000: 55–56).

Diamond (2013) calls this executive function cognitive flexibility and sees task switching as one aspect of mental flexibility. According to her definition, mental flexibility requires the ability to change perspectives. This in turn relies partly on inhibitory control as well as on working memory (Diamond 2013: 149).

Jamadar et al. (2015) define switching as the "ability to dynamically and flexibly adapt changing behavioural context, internal behavioural goals, advance information, or the outcome of a previous response" (Jamadar et al. 2015: 328).

In cognitive psychology, switching and mental flexibility have been investigated by means of different tasks such as the Wisconsin Card sorting test (WCST, (Berg 1948)) and different other task switching paradigms (Jersild 1927). Among the various task switching paradigms is the colour-shape switch task (see e.g. Garbin et al. (2010) or Prior & Gollan (2011)). The colour-shape switch task usually consists of a colour or a shape cue that is shown before the stimulus and indicates the main task that is to be performed (either react to the shape or the colour of the following stimulus. The stimuli consist of two different shapes of two different colours, to which response rules are designed (see e.g. Garbin et al. (2010), López Zunini et al. (2019) or Dong & Liu (2016)).

All task switching paradigms follow the premise that changing a task rule from on trial to another negatively impacts performance in the second trial, since interference is created from the task that was previously relevant. Task switching paradigms compare repeat trials, where a task needs to be carried out on consecutive trials, with switch trials, where the task changes and disrupts successive trials. Switch trials show lower performance compared to repeat trials with regards to accuracy and speed. This is called switching costs, which tend to decrease the longer the response-target interval (RTI) is but never vanish completely. Interestingly, interference from a preceding task as well as the preparation for the following task have an influence on task switching performance, since they activate different cognitive control processes: Dealing with interference that is carried over from the previous task on the one hand, and updating the working memory with the new task on the other hand (Jamadar et al. 2015: 328). Jamadar et al. (2015) used an activation likelihood estimation (ALE) to conduct a meta-analysis of 43 task switching studies that were published before 2014 to determine which brain areas are engaged during task switching. They found that the contrast of switch and repeat trials mostly activates a frontoparietal network including the DLPFC, right VLPFC, left ACC, left SMA, bilateral premotor cortex and inferior parietal cortex. The frontoparietal network showed more activation during switch than during repeat trials. They concluded that while this indicates a central network of executive functions that is active during task switching, the results give no validation that there is a brain region that is only active for task switching and specialized in switch trials (Jamadar et al. 2015: 330). Most task switching paradigms differentiate between proactive control, which refers to cognitive control processes mostly involved in goal setting, and reactive control, which is associated with control processes for dealing with interference that is protracted from inhibiting a response rule (Jamadar et al. 2015: 327). Monsell (2003) explains that the source of the switch costs lie in the task task-set reconfiguration that entails to switch attention between the stimulus attributes and the action rules or the task rules that are associated with the stimulus attributes. This reconfiguration of the task-set allows the selection of the appropriate response and also involves the inhibition of non-relevant elements from the previous task-set (Monsell 2003: 135). Within a task switching paradigm, a switch trial would therefore require the switching of attention as well as inhibitory mechanisms.

Task switching paradigms typically entail single-task blocks where participants must perform the same task subsequently, and mixed-task blocks, where participants must occasionally switch between tasks, but which also contain repeat trials. On this basis, switching and/or mixing costs are calculated. Mixing costs indicate the performance costs on repeat trials in mixed-task blocks and trials in the single-task blocks, while switching costs depict the performance costs on switch trials vs. repeat trials only in the mixed-task block (Braver et al. 2003: 713). More specifically, switching costs reflect the effort of re-activating a decision rule and reconfiguring the necessary response rule for the next response. Remembering the rule associated with the cue is vital (Timmer et al. 2017:246). Furthermore, Braver et al. (2003) assigned sustained and transient cognitive control to task switching. According to this rationale sustained components of cognitive control entail e.g. elevated active maintenance demands that stem from maintaining different task sets on a high activation level or from increased attentional monitoring processes that increases the sensitivity to environmental task switching cues. Contrastingly, transient control processes comprise e.g. internally reconfiguring or updating task goals or connecting cues to the correct stimulus-response (Braver et al. 2003: 714). Within an fMRI study, Braver et al. (2003) applied a task switching paradigm and found increased activation in the left lateral PFC as well as the left superior parietal cortex when contrasting task switch trials with task repeat trials and single task trials. Mixed task blocks both required sustained cognitive control reflected by activation of the right anterior PFC as well as transient cognitive control reflected in left lateral PFC activation and activation of the superior parietal cortex (Braver et al. 2003: 719–722).

Finally, although the executive function of shifting does contain control processes that specifically target switching as a core element, it is likely related to and does include other executive function processes such as inhibiting the prepotent response of the subsequent task and updating the task rule in working memory (Abutalebi &

Green 2008: 564). Shifting has specifically been studied in the context of the bilingual and interpreter advantage (see Sections 5.4 and 6.4) as well as in the context of showing the neural overlap of executive functions and language control (see Section 5.3).

4.3 Working memory updating

Another core executive function revolves around the updating of working memory representations (Miyake et al. 2000: 56). This executive function involves holding information in the mind that is not visually present anymore and manipulating this information (Diamond 2013: 142). The concept of working memory portrays a temporary and limited storage of information while the information is in use (Banich & Compton 2018: 287). Differently put, this means that certain information needs to be kept active in the mind while it is being used for a cognitive task (Baddeley 1983: 311). The updating aspect of this executive function stems from the mechanism that relevant information needs to be manipulated in the working memory instead of just being stored (Miyake et al. 2000: 57). Taking a look back at the basis of executive functions introduced in the beginning of Chapter 4., maintaining a task goal, explains why working memory has such an important part in executive functioning: If information cannot be maintained in the mind short-term and altered thereby, a task goal cannot be maintained, which hampers the ability to plan and execute plans and goals (Banich & Compton 2018: 364). Early lesion studies that showed that in some patients, long-term recall of information remained intact while short-term recall was impaired and vice versa. This indicated that there must be a (partly structural) dissociation between long- and short-term storage of information (Banich & Compton 2018: 288).

Before the concept of working memory was coined, information that could only be retained for a short time and was not stored in long-term memory was called short-term memory and was structurally represented a single component (Shiffrin & Atkinson 1969). Moving away from the concept of a unitary store and focussing on the concept that the information is not just briefly retained but also worked with (Baddeley 2010: 138), different psychological models of working memory emerged: The first model to change how we view working memory was the multicomponent working memory model introduced by Baddeley & Hitch (1974). Their model has been updated several times, but the original model contained three parts: the so called central executive and two supplementary storage systems: the phonological loop and the visuospatial sketchpad. Originally, the central executive was described as a system that coordinates information from the visuospatial sketchpad and the phonological loop and was classified as a limited capacity attention system (Baddeley 1983: 315). However, after criticism in particular from researchers that emphasize the role of attention in working memory models (Cowan 1998, 2008) and criticism of the role of the central executive as simply being a general processing capacity unit, it was developed further on the basis of Norman & Shallice's (1986) model of attentional control (Baddeley 2003: 835). A division was made between two control processes: one process of control relies on the control of behaviour that is guided by cues from the environment and represents an automatic form of control. The other control process is guided by a measure of attentional control in the form of the so called supervisory activating system (SAS) that can take over when automatic control is not enough (Baddeley 2003: 835).

The phonological loop is responsible for auditory content and incorporates a phonological store that holds memory traces for a short time as well as an articulatory rehearsal process that functions as an inner voice. To refresh memory traces, they need to be retrieved and repeated (Baddeley 2003: 829–830). There are different effects that support the model of the phonological loop. For instance, there is the phonological similarity effect: words or letters that sound the same are harder to remember than different sounding items. Furthermore, there is the word length effect: The longer a word is, the harder it is to remember or, the more the memory span decreases. The articulatory suppression effect shows that if remembering a word is interrupted by another non-related word, memory of the target word becomes impaired (Baddeley 1983: 316–317). Evidence from children learning languages has led to the assumption that the phonological loop has developed to facilitate language acquisition. When children learn a language, they need to develop the ability to hear a word they do not know and repeat it. This process is thought to depend on the phonological loop. For patients who show impairments of the phonological loop, acquiring novel words is difficult or impossible. The phonological loop therefore seems to play an important role in long-term phonological learning. The visuo-spatial sketchpad is the counterpart of the phonological loop as it specializes in holding visuo-spatial images in the mind and manipulating them. It also functions as an active storage that is limited in capacity (Baddeley 1983: 319).

A point of critique of Baddely and Hitch’s original model is that it did not make interaction between the phonological and visuospatial system possible. Therefore, another component was proposed to solve this issue: the episodic buffer. The episodic buffer is also assumed to be a temporary storage system of limited capacity that is however able to integrate information from different sources. It acts as a buffer between the phonological loop and the visuospatial sketchpad and long term memory. One of its important tasks is feeding information into and retrieving information from long-term memory. It is also controlled by the central executive (Baddeley 2000: 421).

The multicomponent model by Baddeley and Hitch has been the prevalent theoretical framework for many years. However, while it can account for many aspects of the working memory process, other approaches to working memory have also proven promising in more recent years. So called “state-based models” (D'Esposito & Postle 2015: 117) assume that allocation of attention is the basis of retaining information in working memory for a short period of time. They can be separated into two categories based on the stimuli that are perceived: activated models that mostly deal with symbolic stimuli such as words, letters or numbers, and sensorimotor recruitment models that deal with perceptual stimuli such as colours, sounds and tactiles (D'Esposito & Postle 2015: 117). Approaches such as by Cowan (1988) that fall under the category of activated models highlight different activation states of elements in long-term memory namely in an activated memory that represents short-term memory and a so called focus of attention (Cowan 1988: 180). Over the years, Cowan adapted the model to new theories and insights and also incorporated a working memory element as an embedded feature of long-term memory. According to his newer models, several elements can be more activated than others in the entirety of the long-term memory; as a result of recent perception for example. This is called activated long-term memory and corresponds to short-term memory. The aforementioned focus of attention is another portion of the activated elements in long-term memory and has a limited chunking capacity. Working memory falls back on three components of the model, namely the activated part of memory, the focus of attention as well as a central executive that controls the information that is stored (Cowan 2008: 326–333).

For all activated models, D'Esposito & Postle (2015) summarize that when a piece of information is to be remembered for a short time, its representation in the long-term memory is accessed through a recognition process and kept in an activated state by focussed attention. Especially in Cowans models, the capacity of focussed attention is limited, and attention focus is controlled by a central executive. When an attention shift happens, the items that were in the focus of attention remain in the activated long-term memory which has no capacity limits. However, the activated long-term memory is affected by temporal decay of the information representation (D'Esposito & Postle 2015: 118). Both multicomponent models as well as state-based models share the communality of distinguishing between two components: They have on the one hand a separate system that is responsible for executive control, be it attentional control specifically or a more general approach that governs and influences the memory systems on how to use and manipulate information. On the other hand, both model types have these memory systems that are responsible for storage or retrieval of information. Neurophysiologically, this division is most likely represented by networks communicating in the posterior cortex as well as in the prefrontal cortex (Banich & Compton 2018: 291). More specifically, a neuroimaging review study by Vilberg & Rugg (2008) has found activation during memory retrieval in the left lateral cortex, inferior parietal sulcus as well as the right lateral and medial parietal cortex (praecuneus) (Vilberg & Rugg 2008: 1789). First evidence from parietal involvement in working memory came from a lesion study, where a patient with damage to the left parieto-occipital area exhibited reduced performance in various short term recall tests while his long-term memory seemed unaffected (Shallice & Warrington 1970: 268–269). Event-related fMRI studies such as by Postle et al. (2003) have shown that the retention of working memory representation of stimuli is mediated by posterior regions and not the prefrontal cortex as previously assumed (Postle et al. 2003: 941).

The proposition that information is in fact maintained active across a delay period has been questioned in studies such as by Lewis-Peacock et al. (2012), who used a multivariate pattern analysis of fMRI to show that only item representations in the focus of attention show an active neural trace and not item representations that are kept in activated long-term memory (short-term memory). Despite that items in activated long-term memory did not show activation, they could be remembered across a short delay, accomplished by other neural mechanisms such as transient changes in synaptic potentiation (Lewis-Peacock et al. 2012: 61). These findings show a dissociation in short-term and working memory retention and fit the embedded or state based models of working memory (e.g. Cowan (1998)) rather than Baddeley & Hitch's (1974) multicomponent model.

Apart from posterior parietal regions that seem to mediate storage or reactivation of information, working memory also heavily depends on prefrontal regions that are known to host executive function networks. Prefrontal regions play a part in organizing, monitoring and coordinating behaviour or attention, which fits with the central executive components that are both entailed in multicomponent as well as state-based working memory models. Consequently, prefrontal regions and posterior regions likely work together during working memory usage, the former guiding access to mental representations retained in a posterior network (Banich & Compton 2018: 291).

Concerning testing working memory, it is important to select tasks that require the manipulation of information and not just test holding information in the mind which would test short-term memory. Popular measures are for example complex span tasks (Conway et al. 2005) to measure working memory capacity or the n-back task (Smith & Jonides 1997) that measures working memory updating abilities. The n-back task can consist of different difficulty levels, depending on how many instances need to be remembered. Participants often see a sequence of stimuli and need to decide whether a stimulus matches the N instances before that, depending on the level of recall that should be performed (Jaeggi et al. 2010: 394). As described by Polich (2007), during updating of working memory representation of stimuli, a stimulus is briefly visually processed and its representation is matched with the previous representation in working memory, not unlike mismatch detection. If there is no change in the stimulus attribute, there is no need for updating. However, if a relevant new stimulus is detected, the stimulus representations that are to be kept in working memory need to be updated (Polich 2007: 2129). In the context of the n-back task this can mean that if for example a 2-back-match was detected, the mental representation of the next stimulus that is relevant needs to be updated in working memory, until another match is detected. According to Jaeggi et al. (2010), the n-back task therefore involves working-memory updating processes as well as some inhibitory processes during stimulus matching (Jaeggi et al. 2010: 395).

Working memory processes are involved and interact with other executive functions processes such in inhibition tasks as well as task switching paradigms. Working memory and inhibition processes interact insofar as a task goal must be kept in the mind to identify what needs to be inhibited. This does not mean however, that both concepts cannot be separated in tasks. In a spatial Stroop task that is primarily used to test inhibition, the participant needs to remember two rules. However, the stimuli themselves indicate the location of the response. Consequently, this task places a comparably smaller demand on working memory. The same concept can be applied to the Flanker task, where the arrows tell the participant what to do (Diamond 2013: 143–145). Vice versa, inhibition also supports working memory by inhibiting distractions and controlling attention to keep the mind from wandering (Diamond 2013: 144). Working memory also interacts with the executive function of switching. Diamond (2013) described the involvement of working memory in switching as " […] inhibit (or deactivate) our previous perspective and load into WM (or activate) a different perspective" (Diamond 2013: 149). In task switching paradigms, the participant needs to remember two rules and update the required rule according to the prime that is shown.

Which role working memory updating plays in the context of bilingualism and most importantly simultaneous interpreting is discussed in detail in Sections 5.4 as well as 6.4.

To conclude, describing the executive functions relevant to this dissertation has made clear that although a single executive function can distinguish itself from another by unique mechanisms, there is also considerable overlap between executive functions, the main common factor being to maintain goal representations. As described in the unity and diversity framework by Miyake et al. (2000), the three main executive functions share underlying commonalities and are not completely independent. They particularly describe that some component of inhibition as well as attention control might be present in working memory updating as well as switching tasks (Miyake et al. 2000: 89). In terms of neurophysiology, this chapter has already touched upon where the three main executive functions can be localized and there seems to be overlapping prefrontal-parietal localization of most executive functions.

To consolidate the connection of executive functions and bilingual language processing, the following chapter explains the linguistic involvement of executive functions as well as their neurophysiological overlap in more detail.

5. Executive functions and psycholinguistic models of bilingual processing

A key question in bilingualism and cognitive translation and interpreting studies is how our brain makes enables two languages to interact with one another without causing unwanted interference. Especially in the context of simultaneous interpreting that entails overlapping processes of word identification and word production, this is a crucial point. It has been suggested that executive functions play a role in bilingual language processing to manage this issue. In order to explain this involvement and in view of the interpreting process, language comprehension/reception processes and language production processes needs to be differentiated. Different models for both processes in bilinguals suggest the involvement of general cognitive control in bilingual language processing. At the root of this lies the theory that in bilinguals, access to the mental lexicons is non-selective and both language systems are active simultaneously, which was tested for example in studies with interlingual homographs. Interlingual homographs are words that have the same written form but differ in meaning. Dijkstra & van Heuven (1998) explain that according to a language selective-access to the mental lexicon, interlingual homographs should show faster reaction times (RT) than monolingual control items that are matched in frequency (Dijkstra & van Heuven 1998: 192). Dijkstra et al. (1998) tested this in different lexical decision tasks with English-Dutch interlingual homographs and control words. In Experiment 1, participants had to decide whether a word was an English word or a Dutch word. Results did not indicate significant differences in reaction times for homographs. Experiment 2 replicated Experiment 1, except that now participants were instructed to say “yes” to English words and “no” to Dutch words, creating a higher activation level of Dutch words than in Experiment 1. Results showed a significant difference in reaction time between Dutch and English words, with longer RT for Dutch words with high frequency. It seems that reactions to the English interlingual homographs were strongly inhibited by the Dutch words and participants had difficulty deciding whether a homograph was an English or Dutch word. In this experimental design, a complete selective access and deactivation of the Dutch lexicon would have been helpful. Instead, it seemed that both lexicons were active at the same time and participants had difficulty suppressing words from the Dutch lexicon. These findings speak for a non-selective access to the bilingual lexicon and parallel activation of both language systems (Dijkstra et al. 1998: 52–61). This parallel activation of both language systems is valid for identifying as well as for producing words in two languages. Since simultaneous interpreting heavily taxes both word identification as well as word production, bilingual models for both processes are presented hereinafter.

5.1 Word identification modelling

Based on the language non-selective access hypothesis, Dijkstra & van Heuven's (1998) developed the initial Bilingual Interactive Activation model (BIA model). In the BIA model, words are represented as letter features, letters, words, and language nodes; one language node per spoken language. Initially, letter features, meaning how a letter is constituted, is detected and activates the corresponding letter. Letter units then activate the words containing the letter in the right position. If the word in the position is a mismatch, the activation level of the word is reduced. If the word is a match, it remains activated and feeds activation back to the letters. The language nodes then bundle the activation of all words that are activated in a lexicon and send top-down inhibition to word nodes in the other lexicon. Active word nodes can inhibit other word nodes and this activation goes back and forth between and across nodes. In the end, the word that matches the input best receives the most activation and becomes most active, upon which it transcends a recognition threshold. As already shown in experiments such as by Dijkstra et al. (1998), activation happens more quickly depending on frequency as well as lexical and semantic overlap between words (Dijkstra & van Heuven 1998: 199–200). However, some of the crucial caveats of the BIA model are the missing representation of phonological and semantic features as well as the unspecified representation of interlingual homographs and cognates (Dijkstra & van Heuven 2002: 181). Therefore, Dijkstra & van Heuven expanded the BIA model and created the BIA+ model, that contains phonological and semantic word representations as well as a task schema/decision system that was adapted from Green's (1998) Inhibitory Control Model (ICM). As in the BIA model, lexical orthographic candidates are initially activated simultaneously irrespective of how similar they are to the input and their prevalent activation level. In contrast to the initial BIA model however, activated matches activate their respective phonological and semantic representations. This activation process still remains language non-specific, the similarity of the input word to the lexical, phonological, and semantic representations determines the successful selection of a word. Another difference to the initial BIA model is the integration of a non-linguistic task/decision system and its distinction from the word identification system. The word identification system is affected by linguistic context in lexical, syntactic, or semantic sources such as sentence context. The task/decision system on the other hand is affected by non-linguistic context or so called task schemas such as instructions or task demands (Dijkstra & van Heuven 2002: 187) The most important difference between the BIA+ and the BIA model is that the BIA model assumes top-down inhibitory control from the language nodes while the BIA+ does not support top-down control (Dijkstra & van Heuven 2002: 195). While Dijkstra & van Heuven (2002) propose the task/decision system as a replacement for top-down control by the language nodes and are supposed to engage in the same functions, they do not indicate in their BIA+ model that the task/decision system influences the word recognition system. It seems that the non-linguistic task/decision system is viewed as part of domain-general cognitive control, however, Paap (2018) argues that the word identification system is encapsulated, which makes top-down control from the non-linguistic task/decision system impossible for the BIA+ model. The arrow in the BIA+ model only indicates that the word identification system influences the task/decision system, but not the other way around (Paap 2018: 442). Dijkstra & van Heuven (2002) did consider different possibilities of how the task/decision system could be integrated. They chose the option that the task/decision system does not influence the activation state of words in the identification system but that it can control decisions of e.g. what task these words are used in after the word identification. The BIA + model therefore uses primarily bottom-up processes with the option of top-down processes after word identification. Dijkstra & van Heuven (2002) do however emphasize that there is the possibility that the task/decision system modifies the activation level of words in both languages in the recognition system during the recognition or identification process.

In another extension of the BIA+ model by Dijkstra et al. (2019), the Multilink model that is based on a computational approach, the involvement of the task/decision system in the lexical network is still not conclusively specified (Dijkstra et al. 2019: 666).

Green's (1998) Inhibitory control model (ICM) is one of the models that include an executive component that influences the bilingual lexico-semantic system with a top-down processing approach. It is rather oriented towards production although it also contains mechanisms for word identification. In contrast to models that explicitly target production, the inhibitory control model rather focusses on the language control aspect but does not include processes for articulation (see e.g. Levelt's model further below). The ICM shares some features with other models that touch upon executive functions such as the working memory models elucidated in Section 4.3. Similar to the central executive control system in Cowan's (1998) embedded state-based model that controls and coordinates the allocation of attention, the ICM postulates a control system that lies outside the language control system. It influences language control and is referred to as the supervisory attentional system (SAS). The SAS controls so called task schemas that Green defines as "mental devices or networks that individuals may construct or adapt on the spot in order to achieve a specific task" (Green 1998:69). They can be viewed as a mental representation of the instructions of a task such as "translate into English". If a task has been performed before, corresponding task schemas can be recovered from memory. If the task is new and has not been done before, task schemas can be created spontaneously, which requires more effort, coordination, and control. This application of task schemas that can be automatically retrieved from working memory or must be created under more effort corresponds to early model of executive functions such as by Norman & Shallice (1986) mentioned in Chapter 4. Green's (1998) definition of task schemas leans more on the definition by Norman & Shallice (1986) and differs from the definition of schemas by Dijkstra & van Heuven (2002).

During word reception, the SAS specifies which language is to be used and transmits this to the task schemas. The SAS also controls that conceptual information driven by a goal is transmitted to the lexico-semantic system via the conceptualiser. In the lexico-semantic system, each lexical representation has a lemma that defines its syntactic properties. The lemmas have a language tag that assigns the word to a language. Green (1998) explains that during a translation from L1 into L2, the task schema e.g. for production or translation is initially activated by the SAS. The task schema then increases or inhibits the activation at the lemma level in the lexico-semantic system and active lemmas enter into competition for lemma selection. Consequently, the activation of lemmas with non-target language tags is inhibited. Green (1998) mentions that it is possible for a lexical concept in L1 to also activate a lemma in L2 or vice versa if it shares features with L1, which accounts for the language non-selective access and explains phenomena such as in the lexical decision tasks with interlingual homophones. The activation levels of the lemmas in one language can be skewed depending on language proficiency or conversational context. Therefore, high activation levels of lemmas require more inhibition if they belong to the non-target language. The task schema remains active until the language goal is achieved in the lexico-semantic system and is either inhibited by other schema or if the SAS has altered the task goal (Green 1998: 69–73).

The ICM model was tested by using language switching tests: In their “relative strength hypothesis”, Meuter & Allport (1999) predicted: “The more nearly equal in strength the two languages, the smaller the expected (reverse) asymmetry in the switching costs” (Meuter & Allport 1999: 28). They used a numeral naming task during which L1 dominant participants had to switch between languages following a cue. The results showed longer reaction times for switch trials than for non-switch trials. Furthermore, switching costs[1] were greater when participants had to switch from their non-dominant L2 into their dominant L1 than when they had to switch in the other direction. They explained this effect in accordance with the ICM: If a word in the weaker L2 is to be named, more inhibition is needed to supress the higher activation level of the stronger L1. This inhibition carries over into the switch trial, where the word is to be named in the stronger L1. Therefore, it is more effortful to overcome the amount of inhibition and name the L1, which leads to longer reaction times (Meuter & Allport 1999: 35). Thus, while the BIA, BIA+ and Multilink models

---

[1] Difference in RT between switch and non-switch trials

incorporated executive control rather tentatively, the ICM suggests that there is a cognitive control component that lies outside of the language control system and modulates it top-down.

Other approaches such as by Finkbeiner et al. (2006) have argued that it is not suppression or inhibition of one of the competing lexical representations or task schemas and question the lexical suppression hypothesis that lies at the core of e.g. Green’s ICM. Instead, they argue in favour of the so called “differential activation approach” (Finkbeiner et al. 2006: 1087), in which the target nodes receive additional activation in contrast to the non-target language nodes. They set forth that the additional activation is modulated and regulated by the bilingual speaker’s intention (Finkbeiner et al. 2006: 1087–1088). Finkbeiner et al. (2006) however do not specify how exactly the intentions of a speaker are represented and which circumstances they underlie. From the description of the mechanism of additional activation of target language nodes controlled by intentions, it could be assumed that this is another form of top-down control that requires executive function processes after all, similar to the task schemas in the ICM or BIA/BIA+.

5.2 Language production modelling

Language production models partly rely on mechanisms of language identification models such as the BIA/BIA+ but also incorporate specific processes for speech planning and articulation. Executive functions are integrated in these models at different levels, as will be explained in the following.

Modelling bilingual language production as in simultaneous interpreting has started out with monolingual speech productions models. One of the most popular monolingual models is Levelt’s (1995; 1999; 1998; 1989) model of speech generation that encompasses a conceptualization phase, a formulation phase and an articulation phase. In the conceptualization phase, the speaker basically decides what to say using macroplanning and microplanning processes, generating a preverbal message using working memory and long-term memory (Levelt 1989: 9–10). Levelt (1999) describes these processes as shifting the attention to something the speaker wants to express, monitoring whether what should be said is implemented and laying out a plan to achieve communication goals. The preverbal message is submitted to the formulation phase. Here, lexical concepts in the message activate the syntactic words and their lemmas that belong to them. After lemma selection, morpho-phonological encoding takes place and the morphological and phonological characteristics of the word become available and are matched. During the phonological encoding, the syllables, phrases, and intonation patterns are constituted. The output from this stage is transferred to the articulation phase, where the physical articulatory organs such as lungs, larynx, pharynx, tongue and mouth articulate the message (Levelt 1999: 87–94). Levelt (1995) points out that a network of the cortex, the basal ganglia and the cerebellum are involved in this process (Levelt 1995: 21). In his original speech model from 1989, (Levelt) explicitly talks about executive control and automaticity. He declares that a process is automatic when it is not managed by executive control, since automatic processes are executed without the person being consciously aware of them. Contrary to this, he explains that controlled processing required attentional demands and therefore certain degrees of awareness of the process. In terms of his model, he clarifies that the conceptualization phase involves highly controlled processes, since arranging communicative intentions require much attention and cannot be carried out automatically or by using a pre-existing schema. He explains that monitoring already occurs for internal speech. By this definition, executive monitoring already occurs before the speaker has articulated his message. It could be argued that this planning, goal directed behaviour, attention shifting and monitoring corresponds to the task schemas and top-down control proposed in Green (1998). Levelt (1989) further describes that speakers monitor their speech output in order to correct mistakes or reformulate the message, which also includes monitoring as an executive control function (Levelt 1989: 13–21). More specifically, Levelt (1999) describes that during the articulation process, monitoring as executive control is involved to monitor speech output. He mentions that "[…] the effort of keeping control is more apparent from hesitations, dysfluencies and fresh starts that abound in normal speech" (Levelt 1999: 112). This is another indication of the involvement of executive-functions, since the attention must also be directed to the speech output and goal directed behaviour must be followed.

This modular approach to monolingual speech production is not easily applied to

bilingual speech production since it does not explain how words from one language can fit into syntactic structures from another language and how words in two languages can be jointly activated. Therefore, bilingual models of speech production need to follow a more connectionist and interactive approach (Riehl 2015: 51).
Riehl (2015) argues for an approach that follows spreading activation during speech production as it was implemented in the BIA+ model of word recognition. Words from both languages share concepts that activate semantic representations and semantic neighbours. Parallel to this selection process at the lemma level, activated lemmas trigger their phonological representations and their neighbours. This still happens across languages. The phonological representations that fit best and are most active feed activation back to the lemma level. Consequently, weak candidates from both languages become less and less active, the word that received most activation and is most likely to fit is chosen and articulated (Riehl 2015: 51–52). Riehl (2015) does not specify the conceptual phase or the articulation phrase further but the details from Levelt's (1995; 1999, 1989) model fit around the interactive activation model. It can be argued that the conceptual planning phase in Levelt (1995; 1999, 1989) model initially sets the direction for articulation, upon which the word concepts from the planning phase activate lemmas in both languages and so on. After the words that are to be articulated in the right languages are found they could once more be submitted to the formulation phase including phonological encoding and then transferred to articulation. Riehl (2015) points towards monitoring and inhibitory processes that act as an additional mechanism to keep both languages apart during production. She emphasized that there should be a distinction between a local inhibition that pertains to translation equivalents for a certain concept, or a global inhibition where the activation of a whole language system would be lowered (Riehl 2015: 53–54). In my personal expectation, this is only likely when we are truly talking about lowering the activation of words in a language system and not switching off one language system completely. Riehl (2015) refers to Green's (1998) model that includes the SAS as an executive monitoring system.

These psycholinguistics models for word identification and word production have attempted to gauge the interaction of two languages in one brain and which role executive control functions play in these mechanisms. To test the hypotheses of the models, language switching tasks were often applied (see e.g. Meuter & Allport (1999)). In literature, switching between languages has sometimes been equated to code-switching, however, there are some important distinctions to consider. Code-switching can be defined as "the alternation between the use of one language and another within the same context" (Schwieter & Festman 2023: 114). Yet, there are distinctions made such as by Muysken (2000) who distinguishes between alternation, insertion and congruent lexicalization. During alternation, the speaker alternates between structures from one language to another. During insertion, single words, or entire chunks of a sentence in one language are inserted into the structure of another language. Finally, during congruent lexicalization, words from two different languages can be inserted into a shared grammatical structure (Muysken 2000: 3–5). Another differentiation of code-switching can also be to distinguish between inter-sentential switching and intra-sentential switching. Inter-sentential switching characterises the language switch between two different sentences while intra-sentential switching characterises the switching of languages within a sentence (Kuzyk et al. 2020: 544). Inter-sentential code-switching rather falls into the category of alternation in the terminology of Muysken (2000), while intra-sentential code-switching fit into Muysken's (2000) categories of insertion and congruent lexicalization. Green (1998) has touched upon code-switching in the context of his inhibitory control model. He emphasizes that when switching from one language to another, both languages need to be active to some degree and that code-switching requires a cooperative rather than a competitive relationship between language task schemas (Green 1998: 76). In an extension of the ICM that is tailored to code-switching, the so called control process model of code-switching, Green & Wei (2014) explain that during insertion and alternation, more active language task schemas from one language can give local control to the other language task schemas so that the word from the other language can be inserted. After insertion, the previously more active language receives control again. In case of insertion, it is possible that the inserted word needs to be adapted into the sentence structure, while during alteration, no adaption might be necessary since it could be a whole new sentence. In contrast, during congruent lexicalization or dense code-switching as Green & Wei

(2014) denominate it, the words from both languages that have the highest activation are inserted and adapted into the grammatical structure (Green & Wei 2014: 504). Executive functions are involved here in the form of inhibitory control as set forth in Green's ICM. Which effects frequent switching between languages and code-switching might have on executive functions such as inhibition, attentional shifting and working-memory updating is further discussed in Section 5.4. However, for the context of this dissertation, it needs to be mentioned that simultaneous interpreting and translation do not fit into the categories of insertion, alternation, congruent lexicalization, intra-sentential switching or intra-sentential switching and are therefore not truly code-switching behaviours. While it can be argued that simultaneous interpreting and translation both involve switching from one language to another in the form of translation that requires high activation levels of both languages that are used, especially in the case of simultaneous interpreting, both modalities require the acquisition of translation and interpreting specific skills. Simultaneous interpreting does not involve a code-switch within or between sentences but the overlapping word identification and production processes laid out in this section and the preceding section.

In summary, what becomes clear from the psycholinguistic models that have been presented above is that languages must be jointly activated for bilingual language identification and bilingual language production to be effective. There seems to be little doubt that some form of executive control in language reception and production is involved since all models repeatedly mention processes that are characteristic of domain-general executive functioning at different points. They contribute to the understanding of how languages can be kept apart, and task goals can be reached. However, there seems to be no real consensus among the models concerning the exact point at which executive functions are involved. Especially the concept of monitoring and at which stages it is involved in pre and post articulation does not seem clear cut. The evidence from neurolinguistic studies in Section 5.3 as well as the studies on the bilingual advantage in executive functions in Section 5.4 will further elaborate on the involvement of executive functions in bilingual processing. This will form the basis for the empirical study to see how executive functions are affected during simultaneous interpreting training.

5.3 Neurolinguistic evidence of overlap between executive control and language control

As Sections 5.1 and 5.2 have shown, most psycholinguistic models suggest or do not entirely reject an involvement of executive function in managing more than one language, an assumption that corresponds to the findings of Chapter 3: Various studies have shown how managing multiple languages induces functional and structural changes not only in brain regions related to language, but also in brain regions related to executive function or cognitive control. The question arises whether there is an overlap of executive function networks and language control networks. Several studies have investigated this issue by applying language switching or task switching paradigms as well as by comparing the two tasks. During linguistic switching, participants switch between languages while during non-linguistic task switching, participants switch between tasks. To further explore the notion of overlapping executive functions and language control, several studies from a neuroscientific perspective are discussed in the following.

In an fMRI study, Hernandez et al. (2000) compared a group of bilinguals in single- vs. dual language picture naming. In the mixed-language condition, participants showed increased activity in the DLPFC (Hernandez et al. 2000: 428). Electrophysiological evidence can be found in Khateb et al. (2007) where ERPs of bilinguals were compared in a monolingual task selection and a bilingual language selection in picture naming contexts. They predicted different selection processes during L1 naming in the bilingual and monolingual naming contexts and therefore differing ERP responses. They found diverging ERP patterns at 220ms after stimulus presentation which were interpreted as different L1 naming processes. Additionally, they conducted an estimation by source localization of brain regions and found an increased activation in the left middle frontal-precentral gyri, supramarginal gyri and angular gyri during language selection (Khateb et al. 2007: 208). By applying a task switching paradigm to see whether bilingualism impacts brain areas responsible for the executive function of switching, Garbin et al. (2010) compared monolinguals to bilinguals using MRI. For switch trials, they found increased activation in areas fundamental to the monitoring and inhibition processes in task-switching: right inferior frontal cortex, left IPL and the ACC in monolinguals, and the left inferior frontal cortex and the left striatum as well as reduced switching costs for bilinguals. They concluded that monolinguals and bilinguals use brain networks of executive functions differently and that the increased activity in the left IFG in switch trials for bilinguals clearly shows that there is a relationship between executive control and language control (Garbin et al. 2010: 1277). In another fMRI study, Baene et al. (2015) examined whether there is an overlap in brain activity in linguistic and non-linguistic switching. They found that the lateral and medial PFC (ACC and pre-SMA) and the inferior and superior parietal lobe showed common activation during linguistic and non-linguistic switching (Baene et al. 2015: 1761). Similarly, Coderre et al. (2016) tested bilinguals on a Flanker task with linguistic and non-linguistic distractors and a semantic categorization task while conducting an fMRI scan. They found significant functional overlap in the left IFG for bilinguals for both linguistic as well as executive functions tasks but no functional overlap for monolinguals, concluding that managing multiple languages everyday demands more executive control in bilinguals, which causes the increased functional overlap in the left IFG for bilinguals but not for monolinguals (Coderre et al. 2016: 484). The overlap of executive functions and language control has also been investigated behaviourally, for example by Declerck et al. (2017). They compared language switching costs with task switching costs and implemented the same task format, cues, stimuli and response modalities for task switching and language switching. Results showed no differences between language switching costs and task switching costs as well as significant correlations between task- and language switching costs when a non-linguistic task was used. However, when a linguistic task was used (picture naming), language and task switching costs differed in error rates. They concluded that language switching and task switching share mechanisms only to a certain degree, which is in line with model such as the ICM that assumes a certain degree of involvement of executive function in language control with some language-specific control processes (Declerck et al. 2017: 143–145). In a follow up study, this time using event-related potentials, Declerck et al.

(2021) again compared bilinguals in task switching and language switching. Behaviourally, they did not find differences between language and task switching costs. Electrophysiological results showed no significant differences for cue-locked or target-locked ERPs between language and task switching, also supporting the view that language control and executive functions overlap (Declerck et al. 2021: 640–642).

Not only language switching has shown results pointing towards an involvement of executive functions in language control, but there is also evidence from various translation tasks:

In a positron emission tomography (PET) study, Price et al. (1999) examined the neural underpinning of translation in one direction and switching translation directions. They found increased activity in the ACC and subcortical structures for translation as well as the anterior insula, the cerebellum and SMA, while switching between forward and backward translation primarily engaged the SMG (Price et al. 1999: 2232–2233). Using functional near-infrared spectroscopy (fNIRS), Quaresima et al. (2002) tested bilinguals on translation from their L1 into their L2, vice versa, or in alternating translation direction and found increased oxyhaemoglobin levels in the left inferior frontal cortex that houses the Broca's area (Quaresima et al. 2002: 238). In a silent sentence translation task within an fMRI study that was compared to silent reading, Lehtonen et al. (2005) found increased activation in the left VLPFC as well as the globus pallidus for the translation task (Lehtonen et al. 2005: 609).

As mentioned in Section 4.1, executive functions research often distinguishes between two types of inhibition: interference suppression and response inhibition. Martin-Rhee & Bialystok (2008) examined these two types of inhibition within the context of bilingualism by using the Simon task with an additional delayed answer condition in bilingual and monolingual children. In tasks with bivalent displays, two conflicting dimensions need to be resolved, which requires interference suppression, while in tasks with univalent displays, response inhibition is needed. Their results pointed to a robust advantage for bilingual children in the Simon tasks when a bivalent display was used and therefore, interference suppression had to be applied. When response inhibition needed to be applied, bilinguals did not show an advantage over monolinguals. They argue that the two linguistic systems of bilinguals represent bivalent representations, offering competing response options to the same goals. Bilinguals must focus their attention on the relevant language system and ignore the other system to resolve the conflict (Martin-Rhee & Bialystrok 2008: 91). Similarly, Luk et al. (2010) followed the theoretical case of Bunge et al. (2002a) and examined the neural correlates of interference suppression and response inhibition in monolinguals compared to bilinguals. They used a modified Flanker task with a go/no-go paradigm. Their results showed different neural correlates for monolinguals and bilinguals in the interference suppression condition, but similar networks in the response inhibition condition. They concluded that bilingualism seems to influence brain networks that are used for interference suppression but not for response inhibition. Therefore, although interference suppression and response inhibition are often grouped together as forms of inhibition, they might be cognitively distinct. Furthermore, they observed that the brain areas that were active in bilinguals during interference suppression are congruent with the areas that are a part of the language control network set forth by Abutalebi & Green (2008) (Luk et al. 2010: 355).

The reported regions in these studies indicate an involvement of executive functions in language control and map the relevant regions. Abutalebi & Green (2007; 2008) have given an overview of cortical and subcortical structures that are involved in cognitive control as well as in language control to map a model of executive control and language control. It becomes clear that executive control is a complex function that does not rely on a single system in the brain but is rather implemented by several interacting systems. According to Abutalebi & Green (2007), these systems include predominantly the left prefrontal cortex but also the anterior cingulate cortex, the left inferior parietal lobe and the basal ganglia (Abutalebi & Green 2007: 248). Abutalebi & Green (2008) explain that the prefrontal cortex is engaged in executive functions, more specifically decision making, selecting and inhibiting responses as well as working memory. The PFC cooperates with the ACC which is responsible for conflict detection. Together, the PFC and the ACC identify conflict and trigger a top-down signal. The executive control regions identified by Abutalebi & Green (2007) largely overlap with the executive control regions that Chein & Schneider (2012) associate with their cognitive control network that is most active in the controlled execution stage of skill learning (see Section 2.1). As already mentioned in Section 4.3 in the context of working memory representation maintenance or reactivation, the parietal cortex is also described as part of the executive functions network mapped by Abutalebi & Green (2008). The posterior parietal cortex is directly linked to the prefrontal cortex, which makes it possible to hold information active or on-line. Another network links the PFC to the subcortical structure of the caudate nucleus which belongs to the basal ganglia. Although mostly associated with motor control, Abutalebi & Green (2008) argue that the PFC and basal ganglia interact in such a way, that the PFC maintains a representation while conflict is present and the basal ganglia plays a part in inhibiting unwanted behaviour (Abutalebi & Green 2008: 563–564).

In terms of inhibition, Bunge et al. (2002b) emphasize the strong interconnection of parietal cortex and prefrontal cortex and suggest that both structures are functionally extremely connected and coactivated. By means of a Flanker task they were able to show that the left parietal cortex was activated when there was an increased requirement to maintain a representation of all responses that were possible while the prefrontal cortex was activated when an answer needed to be selected from two competing responses or stimuli types, in this case congruent or incongruent trials (Bunge et al. 2002b: 1566–1569).

As already explained previously in this section, task switching holds a unique and particularly interesting status, since it incorporates not only switching but also inhibition and working memory. Therefore, several studies researching the overlap or involvement of executive functions in language control compare task switching to language switching (see this Chapter). As mentioned in Sections 4.2 and 5.3, Jamadar et al. (2015) and Collette et al. (2005) have identified regions that are most active in studies examining the involvement of executive functions in task switching and have highlighted the involvement of the parietal cortex in task switching.

After describing the neural basis of executive control, Abutalebi & Green (2008) map brain systems that underly language switching based on neural evidence with regions strongly implicated in general executive functions. They delineate a model that includes the left prefrontal cortex, the ACC, the caudate nucleus as well as left and right SMG (Abutalebi & Green 2008: 573). Abutalebi et al. (2012) emphasized the role of the ACC in language control and executive functions. Within the scope of an fMRI study, they asked early bilinguals to perform a language control task in the form of language switching and an executive control task (Flanker). Bilinguals were compared to monolinguals who performed the same executive control task but carried out the language switching task within their only language. They concluded that language control in bilinguals and monolinguals relies among others on the dorsal ACC. Furthermore, compared with monolinguals, functional imaging results revealed that bilinguals used the ACC more efficiently for conflict monitoring, a non-linguistic executive function, indexed by the interference control stimuli of the Flanker task (Abutalebi et al. 2012: 2083–2085).

Green & Abutalebi (2013) updated and specified their original model to also include the thalamus and the putamen as well as the cerebellum in addition to the ACC and pre-SMA, the left and right IFC and left PFC as well as inferior parietal lobe (Green & Abutalebi 2013: 523). In a more recent systematic review, Tao et al. (2021) confirmed the prior established model of overlapping executive functions and language specific-areas by Green & Abutalebi (2013).

This section has examined different models of bilingual processing and mapped the overlap or involvement of executive functions in bilingual language processing to highlight that managing multiple languages and switching between them is a complex interaction of various systems. This is especially important in preparation of defining interpreting as a cognitive task that heavily taxes executive functions and explaining why executive functions can be influenced by extreme forms of language control such as SI. The experimental part of this thesis deals exactly with this question: how are executive functions influenced by simultaneous interpreting training (see Section 6.4)?

As was laid out in Section 3.1, managing two languages in one brain has shown to influence both the function as well as the structure of our brain. Bringing this overlap of executive function and language control together with the knowledge of how bilingualism can shape our brain, the next section dives into how executive functions are affected by the bilingual experience of controlling more than one language and the transfer of expertise from one region to another, before diving into how interpreting affects brain structure and functions as well as executive functions.

5.4 The bilingual advantage hypothesis

As presented in Chapter 2, our brain is capable of changing as a consequence of learning and skill acquisition. Bilingualism has long been discussed as a factor that influences the architecture of our brain (see Chapter 3). Built on the aforementioned models of executive functions, language control and language processing in Chapter 4, the so called bilingual advantage hypothesis emerged. This hypothesis assumes that bilinguals display an advantage in cognitive functions since they have to select the correct word from two language systems which results in enhanced control demands. The two languages should be jointly active but not involuntarily interfere with one another. The bilingual advantage hypothesis is rooted in the construct of simultaneous activation of two language systems in bilinguals and the resulting lexical competition that needs to be resolved to prevent interferences. It is assumed that if bilinguals learn to manage multiple languages from early on and therefore also train their domain-independent executive function due to the overlap in language control and executive control, this should result in long-term advantages beyond the language domain, namely in an advantage in executive functions. This presupposes overlapping neural circuits in prefrontal as well as parietal and subcortical areas for language control and executive functions, as discussed in the preceding Chapter 4 (García-Pentón et al. 2016: 304). Generally, to test the bilingual advantage hypothesis, monolingual and bilingual groups are compared by contrasting behavioural measures such as reaction times and accuracy rates using different experimental tasks. Most frequently, the typical executive functions tests are applied that were mentioned in Chapter 4: the Simon task, Stroop task or Flanker task and the ANT for inhibitory control, the n-back task for working memory updating and e.g. the colour-shape switch task or the Wisconsin Card Sorting Test for switching (van den Noort et al. 2019: 5).

Initially, research literature mainly focussed on the inhibitory control advantage that bilinguals possessed due to their ability to manage more than one language. In their review, Hilchey & Klein (2011) coin the term BICA, bilingual inhibitory control advantage, which they define as follows:

> *"Frequent use of the inhibitory processes involved in language selection in bilinguals will result in more efficient inhibitory processes, which will confer general advantages on non-linguistic interference tasks—that is, those requiring conflict resolution. These advantages will be reflected in reduced interference effects in bilinguals as compared to monolinguals. In other words, bilinguals should show an advantage over monolinguals on trials with response conflict" (Hilchey & Klein 2011: 628).*

These inhibitory control advantages manifested in faster reaction times for incongruent trials in bilinguals relative to monolinguals in tasks such as the Simon task or Flanker task. Evidence that inhibitory processes are indeed required during language processing comes from a study by Emmorey et al. (2008) in which they compared monolinguals, unimodal bilinguals and bimodal bilinguals in a modified Flanker task that included congruent and incongruent trials as well as go-/no-go trials. More specifically, they investigated whether the bilingual advantage in executive functions is in fact encouraged by the need to select only one language while representations from two languages are in conflict for selection. In contrast to unimodal bilinguals, who articulate two languages, bimodal bilinguals articulate a spoken and a sign language and can produce both modalities at the same time and do not experience the constraint to select only one language for production. Results did not show a significant difference in accuracy between the groups, however, there was a significant effect of group on reaction time, with unimodal bilinguals showing significantly faster reaction times than monolinguals and bimodal bilinguals in both congruent and incongruent conditions as well as in go-trials. Bimodal bilinguals did not differ from monolinguals in reaction times. Following these results, the authors suggested that there is not only a bilingual advantage in inhibitory control but also in attention control, monitoring and task switching. They concluded that the bilingual advantage is rooted in competition between two language systems in the same modality (Emmorey et al. 2008: 1203–1204).

The bilingual advantage in executive functions seemed to be modulated by the factor of participant age. Early on, Bialystok et al. (2004) and Bialystok et al. (2005) addressed the bilingual advantage in different age groups and across the lifespan. Building on earlier research with bilingual and monolingual children, Bialystok et al. (2004) conducted studies with bilingual and monolingual older and younger adults that performed different versions of the Simon task. In general, bilinguals showed faster responses in both congruent and incongruent conditions and produced a smaller Simon effect. Interestingly, older bilinguals performed better than older monolinguals and showed a smaller Simon effect. This was taken as support for the hypothesis that lifelong bilingualism decreases age-related decline of inhibitory processing. In another series of studies, Bialystok et al. (2005) tested three bilingual and monolingual age groups on the Simon task. Bilingual children showed an RT advantage in both congruent and incongruent trials over monolingual children in the first and second study. In the third study, monolingual and bilingual young adults between 20 and 30 years were tested. Surprisingly, no difference between monolinguals and bilinguals in both congruent and incongruent trials surfaced. Finally, Bialystok et al. (2005) tested bilingual and monolingual middle-aged adults between the ages of 30-59 and older adults between the ages of 60-80. The results showed that middle-aged adults outperformed older adults and bilinguals outperformed monolinguals in terms of speed. Once again, the bilinguals showed an advantage in both congruent and incongruent trials. Since this phenomenon surfaced in all age groups, the authors suggested that the executive demands specific to the Simon task extend to a more general processing advantage that cannot be limited to inhibition. Furthermore, taken together the studies across the age groups present a U-shaped procession of advantage in executive functions that is apparent in young children, plateaus in young adulthood and enhances processing again in older age. As an explanation for the absence of bilingual processing advantage in young adulthood, the authors suggested that young adults are at their peak performance regarding cognitive processing. They are in control of efficient processing and receive a lot of practice through computer games and other media. The advantage that comes from bilingualism is irrelevant to them and can give no further boost. The authors conclude that in childhood, when executive processes are developing, bilingualism can give a developmental boost as well as in adulthood, when executive processes start to decline, and bilingualism can have a protective function (Bialystok et al. 2005: 107–117).

Coming back to the initial BICA hypothesis, Hilchey & Klein (2011) lay out that the BICA is a rather rarely occurring phenomenon that is bound to very restricted conditions. Furthermore, findings from various studies showed bilinguals outperforming monolinguals in congruent as well as incongruent trials and not only in congruent trials (see e.g. (Costa et al. 2008); Bialystok et al. (2004)). Hilchey & Klein (2011) therefore highlight that instead of just affecting inhibitory control, bilingualism may enhance domain-general executive functions overall. They refer to this phenomenon as BEPA, bilingual executive processing advantage (Hilchey & Klein 2011: 629). According to this rationale, the focus of control lies not on inhibitory processes alone but rather on a comprehensive conflict-monitoring system that functions as general executive control. Since bilinguals must constantly monitor their linguistic representations to select the right one, their monitoring system is more efficient than that of monolinguals (Hilchey & Klein 2011: 654–655). Costa et al. (2008), Costa et al. (2009) as well as Bialystok (2015) supported this view and redirected their focus away from the notion of inhibitory control towards a focus on attention and conflict monitoring. According to Costa et al. (2008), attention has executive components as well, therefore attentional processes are a component of executive control that are involved in selecting the right goal-directed action and involve inhibitory control as well. Furthermore, they argue that executive attention requires both conflict monitoring mechanisms as well as conflict resolution mechanisms. In bilingual speakers, these attentional control mechanisms come to play continuously, which is why hypothetically, the general attention system should become more efficient. They used the Attentional Network Task (ANT) that is widely used to study different attentional networks and tested monolinguals and bilinguals. The results indicated that bilinguals were faster than monolinguals in congruent as well as incongruent trials, and the difference between congruent and incongruent trials was larger for monolinguals than for bilinguals. This was interpreted as evidence that bilinguals possess a more efficient executive control mechanism and are less affected by incongruent trials. Furthermore, it suggests that bilinguals possess a more efficient conflict resolution mechanism than monolinguals. Lastly, results showed that bilinguals had lower switch costs than monolinguals and showed less interference from incongruent trials (Costa et al. 2008: 77–78). Bialystok (2015) sees attention at the core of the bilingual advantage hypothesis and proposes that "[…] the bilingual experience changes the way attention is directed to the environment" (Bialystok 2015: 120). In support of this hypothesis, she references studies such as by Sebastián-Gallés et al. (2012). In their study, silent videos of faces talking in English and French were shown to Spanish or Catalan monolingual and Spanish-Catalan bilingual infants. In the videos, sentences were read in one language first and the language was switched subsequently. Average looking time was analysed. The results showed that only bilingual infants noticed that the language had changed, even if they were unfamiliar with the languages while looking times for monolingual children were at chance levels. This suggests that some dispositions are already developing before children start to use language (Sebastián-Gallés et al. 2012: 994–997).

To sum up, it is not yet clear whether attention control or conflict monitoring as more general concepts of executive functioning are the sole driving force behind managing multiple languages, or whether they rather represent one aspect of executive functioning that coexists next to inhibitory processes and switching processes.

Concerning the executive functions of shifting and working memory updating in bilinguals, studies such as by Prior & Macwhinney (2010) and Hansen et al. (2016) have provided interesting results. Prior & Macwhinney (2010) used a colour-shape switch task to compare the executive function of shifting and monitoring in monolinguals and early bilingual psychology students. They calculated switching costs representing the difference in reaction time between switch and non-switch trials in mixed task blocks, and mixing costs that represent the difference in reaction times in the single-task block and repeat trials in the mixed-task block. Trial type (switch, non-switch) and language group (monolinguals, bilinguals) were used as independent variables while reaction times and accuracy were dependent variables. Results for the switching costs indicated a significant effect of trial type on accuracy and reaction times with non-switch trials being more accurate and receiving faster reaction times compared to switch trials. However, there was no main effect of language group. There was however a significant interaction between trial type and language group for reaction times but not for accuracy. The bilinguals were faster on switch trials than the monolinguals. Therefore, bilinguals showed lower switching costs than monolinguals. Results for the mixing costs did show a significant main effect of trial type (single task, non-switch trials) for reaction times and accuracy with single-task trials being performed quicker and with higher accuracy than non-switch trials. There was however no significant difference between groups and no interaction. The bilingual advantage in this study could therefore only be established in reduced switching costs in bilinguals. Switching costs are said to be created by transient control processes aimed at choosing between two competing tasks. In the colour-shape switch task, these tasks activate the appropriate task goal and change the stimulus-response mapping according to the cue that is shown. The authors suggested that the lifelong practice of the bilingual students in choosing the desired language and coordinating languages as well as in resolving interference between the languages has direct consequences on the executive function of shifting (Prior & Macwhinney 2010: 255–259).

Hansen et al. (2016) examined how different components of working memory develop differently in different age groups of monolingual and bilingual school children by testing them on an n-back task and a reading span task. The n-back task specifically targets executive functioning, especially updating and interference control. With regard to working memory updating, results indicated that only children in the younger age groups (grades 2 and 3) showed a bilingual advantage reflected in correct responses, but not in the older age groups (grades 5 and 8). It seems that emergent bilingualism benefits WM-related processes at an early developmental stage. The authors argue that younger bilingual children but not slightly older children probably are affected most by the transfer effects from bilingualism to executive functions since language interference effects and the demand for executive control are highest at this earlier stage (Hansen et al. 2016: 56–70).

Apart from purely behavioural studies, various studies have investigated the bilingual advantage in executive functions by using fMRI and EEG since in more recent years neuroscientific methods have become more widespread in bilingualism and translation research. The majority of these studies examined structural overlap in executive functions networks in bilinguals or compared activation of said brain networks in bilinguals and monolinguals, as already explained in Section 5.3. Some of these studies describe structural or functional brain differences between monolinguals and bilinguals only and not in combination with behavioural results. García-Pentón et al. (2016) argue that this does not provide a full picture of how bilingual language control and executive control interact. Bialystok (2017) report however that some studies from cognitive neuroscience argue in a different direction, interpreting lacking differences in behavioural performance as beneficial because it allows to look at brain differences with behaviour as a controlled factor. Also, equal behavioural performance but different ERP waveforms between groups could reveal that there is in fact a cognitive difference while the motor response necessary for behavioural results is too coarse-grained to reflect that (Bialystok 2017: 242).

Not relying on structural differences but on connectivity, Grady et al. (2015) compared intrinsic functional connectivity in resting state in monolingual and bilingual older adults. They hypothesized that bilinguals would show increased functional connectivity in executive function networks compared to monolinguals. Results showed stronger intrinsic connectivity in the frontoparietal control network for bilinguals than for monolinguals as well as in the default mode network that is associated with the brain at rest. Other networks not involved in executive functioning did not show any difference. This indicates that older bilinguals’ executive functions still benefit from the modulations caused by bilingualism (Grady et al. 2015: 175–179).

Although the EEG has been used rather infrequently to investigate the bilingual advantage compared to structural methods, there are a few studies that have shown results in favour of the bilingual advantage in executive functions. Kousaie & Phillips (2012) compared performance of bilingual and monolingual young adults on three executive function tasks (Simon, Flanker, Stroop) during EEG recording. They did not find any behavioural differences between the groups. However, they found larger N2 amplitudes for monolinguals than bilinguals in the Stroop task, which seems contrary to the bilingual advantage hypothesis but could indicate that bilinguals were more efficient in their conflict monitoring. The results were however not consistent across all tasks (Kousaie & Phillips 2012: 81). Markiewicz et al. (2023) compared monolinguals and bilinguals in a Flanker task with a with a incongruent/congruent ratio of 25%/75%. For the behavioural results, they found longer response time distribution tails for bilinguals compared to monolinguals, indicating that bilinguals sometimes had longer responses. For the ERP-results, they found significantly increased N2 amplitudes and decreased P3 amplitudes for bilinguals compared to monolinguals. The results were interpreted as evidence for enhanced proactive monitoring in bilinguals (N2 and P3) which resulted in fewer available attentional resources to solve conflict and categorize the stimulus and thus occasionally longer reaction times (Markiewicz et al. 2023: 133–143). Morrison et al. (2019) investigated whether there is a bilingual advantage in working memory using the P3 ERP component. Monolingual and bilingual participants performed an n-back task. Results showed no significant differences in reaction times and accuracy between groups but indicated a smaller P300 amplitude for monolinguals compared to bilinguals, indicating that bilinguals had more cognitive resources available to solve the n-back task (Morrison et al. 2019: 196–197).

Contrary to the positive findings, there are also various studies that did not find a bilingual advantage in executive functions. For example, Paap & Greenberg (2013) compared monolinguals to bilinguals in an antisaccade task, Simon task, Flanker task and colour-shape switch task. They did not find a significant outcome for a bilingual advantage in inhibitory control or task switching and even found a slight disadvantage in one of the Simon tasks (Paap & Greenberg 2013: 242). Testing monolinguals and bilinguals on various task switching paradigms, Hernández et al. (2013) also did not find reduced switching costs for bilinguals compared to monolinguals (Hernández et al. 2013: 271). Bastian et al. (2016) compared monolinguals to bilinguals in different executive functions such as inhibition, conflict monitoring, shifting and generalized cognitive abilities. Rather than evaluating single tasks, the authors built linear mixed-effects models that contained continuous predictors of bilingualism such as age of acquisition, non-L1 language use and proficiency ratio. They did not find an advantage for bilinguals over monolinguals in inhibitory control, monitoring, shifting or generalized cognitive abilities. However, it has to be noted that they did not have a strictly monolingual control group (Bastian et al. 2016: 254). Therefore, within the course of ongoing research, critical voices have continuously debated the validity of the bilingual advantage hypothesis, highlighting the studies that did not find an advantage for bilinguals in executive functions over monolinguals or could not replicate positive results from previous studies. In their review, van den Noort et al. (2019) included studies on the bilingual advantage between 2004 and 2018 and found that while 54.3 % of studies registered a bilingual advantage, 28.3% showed mixed results and 17.4% found evidence against the bilingual advantage. Some critical voices such as Paap et al. (2017) have emphasized the publication bias towards positive results of the bilingual advantage reported by de Bruin et al. (2015) where they reviewed conference abstracts from 1999 to 2012. They reported that studies which found purely results in favour of the bilingual advantage were most likely to be published, followed by studies with mixed results while studies that did not find a bilingual advantage in executive functions for bilinguals were published least (de Bruin et al. 2015: 104). Furthermore, Paap et al. (2015) argues that many studies do not match monolingual and bilingual participants on factors such as socio-economic status, immigrant status or culture which can however influence performance and be confounding variables (Paap et al. 2015: 267–268). In terms of the tasks, Paap et al. (2015) argue that although there are executive function tasks that seem specific to the respective executive function such as the Flanker or the Simon task that target inhibition, two inhibition tasks might not tap into the same kind of inhibition and show little convergence (Paap et al. 2015: 273). In the context of ERP literature, Cespón (2021) also emphasized methodological problems in bilingual advantage studies such as low sample sizes, publication bias and selective reporting of results (Cespón 2021: 2–3). Privitera et al. (2023) have argued that using only fixed-effects analysis methods, as it was done for the majority of bilingual advantage studies, does not do justice to individual differences that can occur within subjects and groups. They recommend using mixed-effects models that are better suited for the data structure in multi-trial tasks since they can include participants and group as factors. They tested this argument by analysing previously collected data from a group of Chinese-English bilingual students performing a Simon task in a linear mixed model and in fixed-effects models that used average reaction times per participant. They found that after including participant and item congruency as random effects all main effects on performance that were identified in the fixed-effects model disappeared. The authors concluded that the heterogenous results in bilingual advantage studies could stem from differences in analysis methods that were applied and also recommended using linear mixed effect models if previous analysis have yielded positive or null effects (Privitera et al. 2023: 2–6).

Although not entirely consistent, a bilingual advantage has surfaced in many studies as described at the beginning of this section. Therefore, Paap et al. (2015) have proposed the following: "A more likely scenario is that bilingual advantages accrue only in very specific circumstances that pair the right set of bilingual experiences with the resonating set of EF measures" (Paap et al. 2015: 275–276). In order to redirect the focus from debating about the existence of the bilingual advantage itself, Woumans & Duyck (2015) have argued that

> *„ […] our field would benefit from moving away from the current yes/no debate. Instead, we should investigate what factors moderate the manifestation of a bilingual advantage, preferably using longitudinal designs and thoroughly normed measures of individual cognitive differences" (Woumans & Duyck 2015: 357).*

Calabria et al. (2018) define three factors that can affect the bilingual language control network: age of L2 acquisition, language proficiency and nature of language use (Calabria et al. 2018: 226). In terms of age of acquisition, they refer to the studies by Berken et al. (2016) and Kousaie et al. (2017) mentioned in Section 3.1. Berken et al. (2016) discovered a link between an early AoA in sequential and simultaneous bilinguals and stronger connectivity between the inferior frontal regions and the right parietal region. In a study using resting-state fMRI, Kousaie et al. (2017) compared simultaneous bilinguals to sequential bilinguals and found that anticorrelations between the VMPFC and the task-positive attention network were stronger for simultaneous bilinguals and that simultaneous bilinguals also showed better interference suppression in a Simon task compared to sequential bilinguals. Regarding L2 proficiency, Calabria et al. (2018) propose that different levels of L2 proficiency lead to different levels of pre-SMA, ACC, left caudate and IFG recruitment, that are all involved in inhibitory control. Finally, they address that the frequency in which bilinguals switch between their languages are likely to affect language control and therefore executive control (Calabria et al. 2018: 226). As a replication and extension of the study by Prior & Gollan (2011), who found a connection between frequent language switching and a bilingual advantage in task switching, Verreyt et al. (2016) also examined how frequency of language switching impacts the bilingual advantage in executive control functions. They administered a Flanker and a Simon task to three groups of participants: unbalanced bilinguals, balanced bilinguals that do not switch between their languages as a routine and balanced bilinguals that frequently switch between their languages in everyday life. All balanced bilinguals had a high level of L2 and acquired both language before the age of six. The results indicate that frequently switching BL display smaller congruency effects than BL that rarely switch between languages at an equal L2 proficiency level (Verreyt et al. 2016: 183–187). Verreyt et al. (2016) argue that an advantage in executive functions requires a frequent activation of lexical representations of both languages as well as the subsequent necessary inhibition of the non-target language, as is the case during language switching. Therefore, the authors place frequent language switching at the core of a bilingual executive functions advantage (Verreyt et al. 2016: 188–189).

Green & Abutalebi (2013) argue in their adaptive control hypothesis that the interactional context is a factor that influences mental control processes and therefore language control and language switching or code-switching. According to the adaptive control hypothesis, control processes that are needed to achieve goals gradually adapt to extraneous demands, that is, different interactional contexts. Not adapting to the interactional context would require an interactional cost. The first interactional context is the single language context that represents the use of one language for a specific environment and another language for another specific environment and does not require frequent language switching. Secondly, the dual-language context entails using two languages with different speakers but not switching between the languages in one utterance but switching between languages within a conversation with different speakers. Finally, the dense code-switching context is represented by frequently switching between languages within one utterance and even the adaption of words from a specific language into another language. Green & Abutalebi (2013) further differentiate between seven control processes that are involved to different parts in the three interactional contexts. The single language context requires goal maintenance and interference control, the latter being further split up into conflict monitoring and interference suppression. In contrast, the dual language context requires salient cue detection, selective response inhibition, task disengagement and task engagement in addition to goal maintenance and interference control. The dense code-switching context however only requires opportunistic planning and is rather neutral towards the other control processes (Green & Abutalebi 2013: 517–519). The single language context requires constant suppression of the non-target language and no reactivation. Contrarily, the dense code-switching context does not require the suppression of the non-target language. Finally, the dual language context requires constant inhibition and re-activation of the non-target language, which can lead to an adaption of the executive functions goal maintenance, conflict monitoring, and interference suppression (Green & Abutalebi 2013: 526–527). Thinking about simultaneous interpreting, it is not straightforward to fit this modality into one of the situational language contexts. However, the dual-language context would fit most to SI, since simultaneous interpreters listen to a speaker in one language and then produce another language. This would also require the constant inhibition and re-activation of the non-target language since both languages need to be kept at a high activation level.

As mentioned in Section 5.2, Green & Wei (2014) explain how code-switching requires the language task schemas to work in a cooperative manner to enable a code-switch. They also integrate the premise of different language contexts from their adaptive control hypothesis into their theory and predict that a dual language context could enhance executive functions in contrast to the single-language context. Assumed that simultaneous interpreting can be categorized as a dual-language context, this would correspond to the bilingual advantage hypothesis as well the interpreter advantage hypothesis (see Section 6.4) (Green & Wei 2014: 506).

Building on Green & Abutalebi's (2013) adaptive control hypothesis and to consolidate the competing results from bilingual advantage studies, Paap (2018) has formulated the so called "controlled dose hypothesis". Similar to the skill acquisition process model by Chein & Schneider (2012) mentioned in Chapter 2, the controlled dose hypothesis assumes that L2 management undergoes different stages from an effortful controlled execution to a more automatic process. The hypothesis relies on the two basic assumptions that (1) a process that requires an excessive use of executive control results in an improvement of executive control ability and that (2) the improvement occurs rapidly but either diffuses because the process is not used anymore or because it has transitioned from the controlled execution stage to the automatic execution stage (Paap 2018: 454). This would indicate that an advantage in executive functions is either connected to the L2 proficiency and only occurs in the beginning of L2 learning or that the processing demands of L1 and L2 switching need to be high enough to tax and train executive functions as well.

Pliatsikas (2020) has proposed a similar approach that is more neuroscientifically grounded by presenting the Dynamic Restructuring model. This model maps the neurocognitive changes occurring as a consequence of bilingualism across three stages. Stage 1 takes effect during initial L2 exposure and learning which demands an increased need to control lexical alternatives and results in a change in a network of parietal, temporal and frontal regions. These regions are implicated in vocabulary, semantic and phonological learning as well as executive functions. These effects result in an increase in grey matter volume in these regions. At Stage 2, immersion and the bilingual experience have increased, which results in a return to baseline volume of brain regions that have previously grown at Stage 1, since only the most efficient brain regions are preserved. This stage is not characterized by fully automated executive functions to control both languages but by a continuous control of lexical alternatives, inhibiting the non-target language and switching between the languages. According to Pliatsikas (2020), while in Stage 1, mainly cortical structures were implicated in executive control and language control, a switch to more subcortical structures takes place in Stage 2. Finally, Stage 3 is characterized by peak efficiency with a renormalization of all prior enhancements as well as automatized language control and all executive functions connected to it. In the context of this model, Pliatsikas (2020) emphasizes that simultaneous interpreting training is an experience that taxes language and executive control so intensely that it may cause the stages 2 and 3 to fuse together (Pliatsikas 2020: 464–466). This approach of a dynamic brain restructuring in reaction to learning, especially passing through different phases, corresponds well to the expansion-partial renormalization hypothesis (see Lövdén et al. (2013)) as well as the skill acquisition model (see Chein & Schneider (2012)) introduced in Section 2.1 in the context of general brain plasticity. In the expansion-partial renormalization hypothesis, training of a mechanic skill in macaque monkeys resulted in an initial increase in grey-matter volume in relevant areas that was followed by a decrease (partial renormalization) in grey-matter after training that was not accompanied by loss of skill. Chein & Schneider's (2012) model of skill acquisition comprises three different phases: formation, controlled execution and automatic execution. Executive control is required most in the controlled execution stage and least during the automatic execution stage. This does not correspond entirely to Pliatsikas' (2020) dynamic restructuring model that places primary engagement of executive function in its initial stage. However, both models assume that training or learning affects the amount of engagement of executive functions and, in the case of the dynamic restructuring model, also structural density.

Grundy et al. (2017a) reviewed MRI and EEG studies to compare the neural correlates in connection with executive functions in monolinguals and bilinguals and come to a similar conclusion as Pliatsikas (2020) in terms of a cortical to subcortical shift with increased executive function proficiency. They summarize that proficient bilinguals show less frontal activation compared to monolinguals who perform the same on executive function tasks, a phenomenon that seems to be reversed in bilingual children. They coin the phrase "bilingual anterior-to-posterior and subcortical shift (BAPSS)" (Grundy et al. 2017a: 188), which describes the pattern of bilinguals recruiting less frontal and executive control such as the ACC and the DLPFC regions and more posterior and subcortical regions such as the basal ganglia in executive functions tasks. Grundy et al. (2017a) describe this pattern as an indicator of increased efficiency (Grundy et al. 2017a: 188). Three studies especially fit this observation:

Mohades et al. (2014) compared early simultaneous bilingual children to early sequential bilingual children and monolingual children in a Simon and Stroop task using fMRI. They found that both bilingual groups showed higher accuracy in both tasks but at the expense of reaction times. Furthermore, there was increased activation in incongruent trials in the STG, bilateral cingulate gyrus, middle frontal gyrus and right caudate nucleus in bilinguals compared to monolinguals and increased activation in the caudate nucleus and posterior cingulate in simultaneous bilingual children compared to sequential bilingual children (Mohades et al. 2014: 632–635). Studying ageing bilinguals, Gold et al. (2013) compared older adult bilinguals and younger adult bilinguals to older adult and young adult monolinguals in a task switching paradigm using fMRI. In the first experiment, older bilinguals outperformed older monolinguals in switching performance. In the second experiment, bilingual older adults and bilingual younger adults performed better than younger and older monolinguals while showing less activation in left lateral frontal and cingulate cortex. Since bilingual older adults outperformed monolingual older adults, this was taken as an indicator for increased neural efficiency in executive function processing in bilinguals (Gold et al. 2013: 392–393). Ansaldo et al. (2015) took these results one step further by comparing elderly monolinguals and bilinguals on a Simon task and a Stroop task using fMRI. In the Stroop task, bilinguals outperformed monolinguals in the incongruent condition. In the Simon task, both groups performed equivalently in terms of reaction time and accuracy. However, monolinguals activated the right middle frontal gyrus while solving incongruent trials, whereas bilinguals rather activated left inferior parietal regions. Normally, there is a shift from posterior to anterior regions in aging brains, which could only be observed in monolinguals in this study (Ansaldo et al. 2015: 462–465).

Taken together, there is recent neuroimaging evidence that the bilingual brain not only changes when it is presented with the challenge of managing multiple languages, but it also renormalizes these changes structurally or shifts tasks to different brain areas while keeping up the efficiency as shown in studies that corroborated the structural neuroimaging findings with behavioural results.

This final section of Chapter 5 has now merged all aspects of neuroplasticity, bilingualism and executive functions so far. It has become clear that although managing multiple languages does induce changes in the brain, it would be an oversimplification to only speak of an advantage in executive functions. Whether executive functions change as a result of managing multiple languages seems to depend on different aspects of the bilingual experience such as L2 proficiency, how often a person uses both languages and switches between the languages as well as processing demands of the linguistic task at hand. In the end, this chapter has also made clear that there is a possibility that a change in executive functions induced by an increase in language control is not stagnant but renormalizes when the language or linguistic task is automized.

The search for what modulates linguistic-induced changes in executive functions has put simultaneous interpreting in the spotlight since it requires managing multiple languages under potentially stressful conditions and time pressure. As the next chapter will further elaborate, interpreting is a cognitive task with high processing demands that requires a high degree of language and professional competence, executive control and attention. How simultaneous interpreting training can affect executive functions and why this dissertation can be an important contribution in this research area will also be discussed further.

6. Interpreting as a cognitive task

Within the bilingual advantage debate, the search for cognitive advantages in bilinguals has put the spotlight on interpreting as a cognitive task that requires high levels of both language control as well as cognitive control. This chapter therefore highlights the research pillars of cognitive interpreting studies and answers questions such as: What characterises simultaneous interpreting? Which processes are at work? Which interpreting strategies are applied and where might executive functions be involved? Furthermore, the following sections will explain simultaneous interpreting as a profession, but also how simultaneous interpreting training changes the brain and how executive functions are affected. Finally, a look at models of interpreting competence will emphasize the importance of training and will draw the connection between interpreting training and neuroplasticity.

As pointed out in Pöchhacker (2016), a differentiation between the two main working modes of interpreting, consecutive interpreting and simultaneous interpreting, was made from the 1920s onwards (Pöchhacker 2016: 18). While consecutive interpreting is characterized by a target text that is produced after a source text or speech, simultaneous interpreting (SI) requires the interpreter to listen and speak concurrently, resulting in a source message that is rendered instantaneously (Ahrens 2017: 446). The simultaneity of the process of SI that involves parallel processing of bilingual output and input is seen as an exceptional ability, which is why most early research has focussed on the simultaneous interpreting mode, advancing conference interpreters into the research focus (Ahrens 2017: 449). Fabbro & Darò (1995) highlighted the simultaneity of processes during simultaneous interpreting and emphasized that interpreters need to listen to the message in the source language, translate the message into the target language and formulate the message in the target language all at the same time while adhering to severe time constraints (Fabbro & Darò 1995: 310). Seeber (2011) argues however that simultaneous interpreting rather represents a temporal overlap of language comprehension and language production and not exactly simultaneous processes (Seeber 2011: 185). Frauenfelder & Schriefers (1997) explain that language comprehension and language production lie at the core of the interpreting process. They emphasize how complex it is to establish an accurate model of simultaneous interpreting, since language comprehension and language production are constantly overlapping during interpreting and are not easy to research in the monolingual context, let alone in the bilingual interpreting context (Frauenfelder & Schriefers 1997: 73–74).

This overlap of word recognition and production is characteristic of SI and therefore differentiates SI from other bilingual modalities such as bilingual speech and code-switching as well as written translation. As laid out in Sections 5.1 and 5.2, there are good reasons to assume that both languages are active during bilingual language reception and production (Dijkstra et al. 1998; Paradis 1994), that executive functions are implicated in this process to guide and support the word identification and production process (Green 1998; Abutalebi & Green 2007; Green & Abutalebi 2013), and that therefore, bilinguals may have increased control demands over their languages at least in certain circumstances of their bilingual profile (depending on AoA, proficiency, processing demands) (see Hilchey & Klein 2011, García-Pentón et al. 2016). As Christoffels & de Groot (2005) explain, SI and written translation do share some basic principles, the most evident being that both tasks tap into similar processes of bilingual language processing to achieve communication for a third party. Translators and interpreters are not allowed to muddle their personal beliefs into the context of the message, and they need to self-monitor during language and text production. However, input and output modes of SI and written translation are fundamentally different. While written translation relies on a visual and written mode, SI is usually characterized by the auditory and verbal mode[2]. Furthermore, the input rate during SI is determined by the speaker of the source text and is usually comparable to the input rate of normal speech (about 100-200 words per minute). Consequently, simultaneous interpreters need to produce output directly, with a speaking rate of about 4.000 words in a 30 minute speech. In contrast to that, the source text in written translation is fixed and the translator can go back to the source text to reread certain passages. This is not possible during SI, and, together with the

---

[2] There are also hybrid forms such as sight translation, where a written source text must be interpreted orally into a target output

high speaking rate, SI creates an urgency in the action that is not given during written translation. Therefore, only simultaneous interpreting is characterized by the near simultaneity of speech reception/comprehension processes and speech production processes (Christoffels & de Groot 2005: 455–456). These overlapping processes performed under time pressure create a higher level of control demands than other bilingual modalities (van de Putte et al. 2018: 244). As Riccardi (2005) summarizes:

> *"Compared to monolingual communication, SI is an unnatural form of communication, whose main peculiarity – in addition to its bilingual nature, common to all oral and written translation forms – is given by the time pressure under which it is carried out" (Riccardi 2005: 756).*

The overlap of source text or language input and target text or language output is therefore the most prominent feature of simultaneous interpreting and has been studied since the 1950s as so called ear-voice span or décalage. The time lag between the speaker's input and output is suggested to reflect cognitive processing of the interpreter and is affected by different factors such as the speaker delivery rate, text type, applied interpreting strategies and interpreting competence (Timarová et al. 2011: 121–123). Studies such as by Treisman (1965) and Lee (2002) have shown that factors such as longer sentence length, interpreting in the direction of the non-native language or task difficulty (shadowing vs. SI) that require more cognitive effort lead to longer ear-voice spans and other factors such as longer between-sentence pauses decreased the ear-voice span, possibly because information could be properly processed, reducing cognitive load.

Over time, many different models have been developed to explain the elaborate interpreting process and more and more important subprocesses have been added to the picture, a relevant selection of which will be presented in the following section.

6.1 Cognitive models of simultaneous interpreting

The earliest interpreting models were product-oriented and based on observations of interpreters, but also process-oriented when based on retrospections (Ahrens 2017: 448). One of these early models based on an observational approach was established within the framework of the École Supérieure d'Interprêtes et

Traducteures (ESIT) by Danica Seleskovitch and Marianne Lederer: By studying conference interpreters, they formulated three phases of the interpreting process which they denominated Théorie du sens. The first phase is characterized by understanding the sense of an uttered phrase through the use of different cognitive complements, which can be understood as different forms of background knowledge about the world. This is followed by the deverbalization phase, where the verbal forms of the utterances disappear and are reformulated according to their true meaning in the target language; these are finally expressed in the last phase (Seleskovitch & Lederer 1989: 36–42). However, although this was one of the first approaches of using empirical data to study the cognitive processes of interpreting and although this theory was formulated by upholding ecological validity, the lack of experimental design and experimental hypothesis testing neglected the observation and interpretation of important cognitive sub-processes and executive functions. Furthermore, results from observing conference interpreters, although a promising method at the time, can be subjective and are not free of biases (Garcia 2019: 16).

Chernov (1979) proposed a model that accommodates the interpreter's cognitive multitasking on the basis of redundancy of languages. According to this reasoning, redundancy of languages enables probability prediction on a semantic level when a message is received in a source language. Furthermore, an anticipatory synthesis takes place, while the message in the target language is generated (Chernov 1979: 278). Anticipation processes interact with the interpreter's memory store of general world knowledge and with the context of the situation, making them an integral part of interpreting as a cognitively demanding task (Chernov 1979: 293).

Moser's (1978) flow also puts emphasis on anticipatory processes as well as memory storage. Initially, incoming acoustic input is analysed for its features by using the phonological rules that are stored in long term memory. When the syllables are recognized, they are transformed into words of the source language with the help of semantic and syntactic cues. Next, strings of processed words are stored temporarily in the so called Generated Abstract Memory (GAM), a storage that is similar to short-term memory, creating a feedback loop that makes it possible for newly incoming information to be further processed. Information in the GAM is chunked into units of meaningful phrases of different sizes. The GAM has a limited number of these chunks it can hold, which is why less and less capacity is available for new storage the more capacity is used to chunk larger units. The GAM works together with language concepts in long-term memory to find a conceptual base for the processed information. Once conceptual relations are activated, the meaning of the words can be understood, and the corresponding target language nodes can be activated. To deal with the processing capacity of the GAM, the interpreter uses prediction or anticipation processes. When the interpreter is aware of what is coming next, many processing steps until the activation of target language equivalents are omitted (Moser 1978: 354–360). The interaction between memory systems as well as between memory systems and anticipatory processes was an important asset for later models.

In the 1980s and 1990s, methods and paradigms from cognitive psychology and psycholinguistics as well as a need for more scientifically based research greatly impacted research in simultaneous interpreting (Gile 2015: 45–47). Based on findings from lesion studies in aphasia patients, Paradis (1994) formulated a neurolinguistic theory of simultaneous interpreting. It is based on the activation threshold hypothesis, assuming that items in the multilingual lexicon have different activation thresholds. The more often an item is activated, the lower its activation threshold becomes and the less stimulation it needs to be activated. When an item is selected for activation, all competing items are inhibited by raising their activation thresholds and decreasing their activation. During simultaneous interpreting, items from both languages are activated but not to the same degree. Therefore, Paradis (1994) assumes that interpreters who switch frequently between languages have probably developed lower activation thresholds for the language that is not selected, so they can alter between comprehension and production fluently (Paradis 1994: 320–321). In a complex flow model, Paradis (1994) highlights the overlapping cognitive processes of encoding and decoding that take place during simultaneous interpreting and also points out the differential involvement of the left and right hemisphere in language processing. Ultimately, he also emphasizes that implicit linguistic competence and explicit metalinguistic knowledge are highly involved in simultaneous interpreting (Paradis 1994: 324–332). Setton’s (1999) model is based on theories about how perception, cognition and action interact in speech processing. His model contains representations of speech input, word recognition, decoding and assembling, world-, situational- and linguistic knowledge and working memory as well representations for formulation and the articulation processes. Interestingly, Setton (1999) specifically includes an executive that not only coordinates tasks but is also involved in meaning assembly and controls production (Setton 1999: 63–92).

A first attempt at a neuroanatomical model of simultaneous interpreting was constructed by Fabbro (1999). Based on a series of experimental studies on the neuropsychological aspects of simultaneous interpreting, Fabbro (1999) set forth a model of simultaneous interpreting that involves the main functional components that are implicated in the SI process and also highlights the involvement of both hemispheres during simultaneous interpreting. The author describes that the left hemisphere is mostly implicated in recognizing the message in the SL and translating and articulating the message in the TL. The right hemisphere is rather implicated in emotional-attentive tasks and aspects of nonverbal communication The model shows the involvement of the bilateral primary and secondary auditory cortex, left-hemispheric subcortical structures, left temporo-parietal cortex, right temporo-parietal areas, left inferior anterior frontal cortex, left and right premotor areas, left supplementary motor area and Broca’s area. Furthermore, the model shows an involvement of the cingulate gyrus bilaterally. Although some of these brain regions are language and motor function specific, especially the mentioning of the left inferior frontal cortex, the left SMA and, although not further specified, the cingulate cortex and subcortical structures overlaps with the model of executive functions and language control of Green & Abutalebi (2013) described in Section 5.3. (Fabbro 1999: 205–206).

In Garcia (2019), the author contrasts his neuroarchitectural translation model (NTM) that is not interpreting specific but depicts the organization of both translation and interpreting systems with Fabbro’s neurocognitive model of simultaneous interpreting. Garcia’s model is mostly based on lesion studies and leans on other findings about the bilingual brain. Its most important neuroanatomic regions for translation and interpreting incorporate left hemispheric perisylvian and frontostratial regions. Garcia (2019) acknowledges that although in contrast to the NTM, Fabbro's (1999) model does not capture the aspects of conceptually mediated and form-level connections, it represents separate routes for forward and backward translation and is generally based on empirical methods. Therefore, it can be further tested (Garcia 2019: 126–128). What is striking about both models is that they do not incorporate specifically how executive functions play into the interpreting process, although models of language identification as well as production have partly incorporated executive functions (see Sections 5.1 and 5.2) and various studies have shown an overlap between executive functions and language processing (see Section 5.3).

Daniel Gile's research focus lay primarily on effort during interpreting. He based the foundations of his effort model of interpreting on findings about attention and mental capacity from cognitive psychologists like Kahneman (1973). Kahneman's capacity theory assumed that "[…] there is a general limit on man's capacity to perform mental work" (Kahneman 1973: 8). Proceeding from this, Gile explained that there is a distinction between non-automatic mental operations that require attention or processing capacity and automatic mental operations that do not. Mental operations that are non-automatic use up processing capacity from a limited repository. Consequently, non-automatic operations were thought to be involved in interpreting, such as identifying new or unknown stimuli, identifying known stimuli that are presented under difficult conditions, storing and altering information in memory or preventing incorrect responses (Gile 2009: 159–160). Gile proposed a model that encapsulated the most effortful operations during interpreting and depicted them in the formula SI = L + P + M + C.

The effort model therefore consists of four central effort components:

- The listening as analysis effort (L) that comprises all operation that are directed towards comprehension, starting from analysing soundwaves of SL utterances over identifying words to finding the meaning of the SL speech.
- The production effort (P) that is classified by converting the mental representation of the utterance into a speech plan and then carrying out that speech plan as well as monitoring as self-correcting the output.

- The memory effort (M) which mainly revolves around the effortful operation of the short-term memory.
- The coordination effort (C) that is reflected by the resources necessary for the coordination of the listening, production, and memory effort

According to Gile, more than one and even all three core efforts can be active at the same time, however, each effort has a limited processing capacity. Therefore, in order for interpreting to unfold fluently, the processing capacity requirements for L, P and M should not overstep the respective as well as the overall available processing capacity of the interpreter. The interpreter must therefore learn to coordinate the resources to the important processing capacities at the right time (Gile 2009: 160–170).

Also occupied with the resource capacity during interpreting, Seeber (2007) addresses the overlap of the effort processes identified by Gile and establishes a two-dimensional model of the cognitive resource footprint for shadowing, sight translation and simultaneous interpreting. In the individual cognitive resource footprints, the processing stage, modality and the sequence of the task or processes at hand can be visualized. An interference score can be calculated that describes the interference between resource-demanding perceptual tasks such as listening or reading and cognitive tasks such as working memory involvement when they are executed simultaneously (Seeber 2007: 1382–1384). However, Seeber (2017) concedes that the assumptions for the cognitive resource footprints for shadowing, sight translation and SI are only valid under laboratory settings and do not apply to real-life interpreting situations (Seeber 2017: 467).

Already in 1969, David Gerver highlighted the importance of monitoring and memory in simultaneous interpreting. After comparing SI during shadowing and interpreting in a study, he suggested that interpreters store input while they are translating the SL utterance and monitor the TL output. He concludes that attention during interpreting can be split between receiving input, translating, and monitoring. However, if a certain capacity is overstepped, allocating attention towards input or output suffers (Gerver 2002: 65–66). Gerver subsequently proposed a flow model that incorporated different kinds of memory systems. Among these were a short-term working memory buffer that ensures the ongoing translation process while new information comes in, a short-term memory output buffer that corrects and monitors the output that is to be produced, as well as a long-term memory that stores language information. Building on research suggesting that attention can be split between different tasks, he also emphasizes that control processes are involved that monitor and determine how attention is distributed during the interpreting task. He concluded that with this divided attention, the interpreter can handle different tasks under normal condition but that, in the case of occurring difficulties, attention must be refocussed on decoding and encoding. Resulting from this, monitoring of input and output can be compromised (Gerver 1976: 191–193).

At this point, it is important to take a closer look at the term "monitoring" since it occurs in the executive function literature reviewed in Section 4.1, in the context of bilingual word identification in Section 5.1, in the context of language production in Section 5.2 as well as in the context of interpreting presented in the current chapter. However, not all notions of monitoring refer to the same concept, which needs to be distinguished here, although there are also overlapping aspects. In the executive function literature, monitoring is mostly understood as conflict monitoring and is understood as a process that detects conflict and then initiates actions to resolve this conflict for example through interference suppression or inhibition (see Botvinick et al. (2001)). In Section 5.1, conflict monitoring is taken up indirectly in the context of word identification, where especially the ICM by Green (1998) includes executive functioning in the form of the supervisory attentional system. When a word is heard, the SAS already specifies which languages is to be used and sends this via the conceptualiser to the lexico-semantic system where the information influences the activation levels of word representations. When it comes to translation or answering in another language, inhibition is presented as the mechanism in focus. Explaining language switching costs, Meuter & Allport (1999) assume that when switching from L1 into the presumably weaker L2, more inhibition is necessary to inhibit the activation of stronger L1 representations. This could also be rephrased as conflict detection. So, a lot of conflict is detected when switching from L1 into L2, which requires inhibition of activated L1 representations. The prolonged reaction times during switching back from L2 into L1 stem from the increased inhibition of the L1 representations that needs to be overcome. Monitoring as part of executive functioning clearly would take place at the word activation level. In the context of word production in Section 5.2, Levelt (1989) does not refer back to the concept of conflict monitoring but rather to the concept of error monitoring by monitoring internal speech at the early conceptual phase and monitoring speech output during articulation. He placed monitoring in the realm of executive functions but since it is a monolingual speech production model, does not take conflict monitoring for conflict with other languages into account. Monitoring as it was mentioned in the context in translation studies and here in interpreting studies has a similar character. Translators and interpreters have been assumed to use different routes to arrive at their translation: vertical translation requires the translator or interpreters to fully understand the meaning of the source language text or utterance and happens via a meaning based production of the target word. Horizontal translation on the other hand is based on replacing linguistic structures of the source language with the corresponding structures of the target language (de Groot 1997: 30). Vertical translation has often also been referred to as the meaning-based or semantically-mediated route and horizontal translation has been referred to as the word representation-based route or transcoding (see Schaeffer & Carl (2015) 2015 or Amos & Pickering (forthcoming)).

Referring back to Toury (1995), Tirkkonen-Condit (2005) explains that monitoring during translation alerts the translator about a problem that occurs in the outcome if a literal translation is followed. Literal translation is seen as the default method of translation that is interrupted when conflict is detected. In this case, monitoring seems indeed to be understood as conflict monitoring in the sense that there is conflict due to a missing literal translation equivalent that can be resorted to. Tirkkonen-Condit (2005) mentions that there could be a difference in monitoring between novice and expert translators since novices would tend to resort to literal translation more than experts. However, she acknowledges that literal translation can be observed in experts as well (Tirkkonen-Condit 2005: 407–408). In line with Tirkkonen-Condit’s monitoring model, Carl & Dragsted (2017) assume that during written translation, text production and comprehension processes occur in parallel, which is described as the default procedure by Tirkkonen-Condit (2005). While the text is produced, the eyes gather the required information to continue the text production, which however goes against the mind-eye-hypothesis (see Just & Carpenter (1980). The monitor interrupts the default procedure when translation problems occur. Carl & Dragsted (2017) collected key-logging and eye tracking data in a copy task and a translation task. They found that participants showed long instances of consequent text production that was not stopped by the monitor in both tasks. However, during translation, a phrase was sometimes typed before the translator knew the exact translation of the rest of the phrase. The authors suggested that meaning in the source text emerges and consolidates while the translation progresses. When translation problems were encountered, signalled by the monitor, these problems were triggered by production problems and not by comprehension problems in the ST (Carl & Dragsted 2017: 6–28).

In terms of simultaneous interpreting, Amos & Pickering (forthcoming) summarize that during SI, the meaning of the input utterance needs to be fully understood to use the meaning-based route and that the word representation or transcoding route is seen as inferior since it is mostly used by interpreting students that heavily rely on translation equivalents and must yet learn to pay attention to the meaning of the input utterance. They emphasize however that more recent views suggest that both routes may be applied during interpreting also by experienced interpreters, but that interpreting mostly occurs via the semantically-mediated route. Whether conflict monitoring then sets of a trigger during non-semantically mediated transcoding, indicating that translation needs to be redirected from pure translation equivalent activation to meaning-based translation is not specified in Amos & Pickering (forthcoming). In a study where she investigated the influence of audiovisual input on self-monitoring during interpreting, Gieshoff (2017) calculated the number of cognate translations when the interpreters worked with or without visual input of lip movements. She explained that interpreters are encouraged to avoid especially low frequency cognates that are uncommon in the target language which requires interpreters to monitor their output more closely when coming across cognates. Inhibiting those cognates can therefore be classified as successful self-monitoring.

She found that without visual input, cognate translations in interpreters increased and interpreted this as an indication for less effective self-monitoring (Gieshoff 2017: 317–323). Gerver (1976) mentions the monitoring of input and output in simultaneous interpreting but does not specify whether this refers to conflict monitoring of competing languages or simply monitoring which language is the input language and whether the translation output is correct during articulation. Input monitoring and output monitoring can still be a part of executive control since it requires the allocation of attention to these processes, however, they do not seem to correspond to conflict monitoring that is connected to interference suppression that lie on a "deeper" level of language processing. Maybe at this point, a model of the interpreting process would be helpful that included different layers of processing, surface layers and deeper layers. Since these monitoring processes are not specified anywhere in a comprehensive manner, I would suggest to at least differentiate between "paying attention to" or monitoring speech input during word identification, monitoring conflict from different competing language representations or translation options in the mental lexicon and inhibiting or suppressing interfering word representations that are activated during the internal speech formulation process, and lastly "paying attention to" or monitoring speech output during articulation.

Most of these models have highlighted specific aspects of the interpreting process that are not language specific such as monitoring and split-attention (Gerver 2002), working memory (Moser 1978), anticipation (Chernov 1979), cognitive effort (Gile 2009; Seeber 2007,, 2017) or activation and inhibition (Paradis 1994) with the later models aiming at a more comprehensive view (Setton 1999; Paradis 1994; Fabbro 1999). However, apart from Setton's model, there is no specific mention of executive functions, although monitoring, working memory and activation/inhibition clearly fall into the realm of executive control (see Chapter 4). This seems necessary however, since simultaneous interpreting is clearly a modality of bilingualism and therefore must require the involvement of executive functions to manage the enhanced control demands.

This section has summarized how interpreting studies over time have shed lights on different aspects of the cognitive mechanisms of interpreting and have evolved from an observational methodology towards more controlled experiments and more comprehensive models that also incorporate empirically-based neuroarchitectural features. Executive functions are mostly not mentioned explicitly and, if so, not in a detailed manner (Gerver (2002); Fabbro (1999); Setton (1999); Chabasse (2009)). However, especially aspects such as working memory, inhibition and activation that are part of the aforementioned models clearly belong to executive functions. Together with Fabbro's neurocognitive model of interpreting that mentions brain areas that are also implicated in executive control (see Green & Abutalebi (2013)), this section shows that executive functions do play a role in models of interpreting, but they have not been given specific and comprehensive attention in pertinent models.

As a result of the previous chapters, SI has been positioned as an extremely demanding cognitive task by presenting models of the interpreting process and highlighting the challenges interpreters face when managing the overlapping processes of listening and speaking such as cognitive multitasking (Chernov 1979), changing activation thresholds of both languages in use (Paradis 1994), capacity limits and associated cognitive effort (Gile 2009) or split attention (Gerver 2002). Therefore, the question arises what the interpreters must do to deal with these challenges. As Riccardi (2005) points out: "Conference interpreting is a goal-oriented activity directed towards the faithful reproduction of the source text in the TL under given circumstances" (Riccardi 2005: 764). Riccardi (2005) explains that since interpreters need to achieve this goal, they must engage in decision-making and problem-solving activities. She argues that since decisions need to be made consciously in the context of SI which is why SI is based on strategic behaviour (Riccardi 2005: 764). SI novices must deliberately train to learn how to solve interpreting-specific problems that can be related for example to syntactically reconstructing the source message so it matches the target language that is uttered. The SI trainees therefore develop strategies through training to solve the problems encountered during the SI process and increase their SI competence to finally become experts (Moser-Mercer et al.

2000: 110). It can therefore be argued that learning specific strategies to successfully interpret an utterance and SI competence acquisition go hand in hand. In the context of this dissertation, competence acquisition is a central topic since an increase in SI competence has been shown to be accompanied by enhanced executive functions (see e.g. Yudes et al. 2011, Dong & Liu 2016, Dong & Zhong (2017) in Section 6.4). The following chapter therefore discusses the notion of interpreting competence, differentiates it from related terms such as expertise and skill, and elaborates on strategies that need to be acquired during SI training to foster SI competence acquisition.

6.2 Modelling interpreting competence

When talking about interpreting competence, first of all, it is important to distinguish between the notion of competence and the notion of expertise. Sternberg (2005) defines the development of competence as "[…] the ongoing process of the acquisition and consolidation of a set of skills needed for performance in one or more life domains at the journeyman-level or above" (Sternberg 2005: 15). On the other hand, Sternberg (2005) defines the development of expertise as "[…] the ongoing process of the acquisition and consolidation of a set of skills needed for a high level of mastery in one or more domains of life performance" (Sternberg 2005: 15). He consequently distinguishes between being able to perform a skill competently and being able to perform a skill at expert level. In translation process research, expertise and competence have often been used as synonyms in the past. However, as argued by Tiselius & Hild (2017), competence and expertise are different from each other and should also be investigated as two different concepts. They define competence as "[…] a set of different capacities and skills necessary for completing a translation or interpreting task" (Tiselius & Hild 2017: 426). They see competence as a prerequisite for the development of expertise and define expertise as "[…] the mastery of outstanding skills by an expert, a mastery that is only achieved after many years of goal-focused work and deliberate practice" (Tiselius & Hild 2017: 426). One could therefore say that they applied the general definitions by Sternberg (2005) to the translation and interpreting studies context. In an attempt to investigate expert performance in the domain of translation and interpreting, Ericsson (2000) discusses the importance of deliberate practice for expert performance. Analogous to other skill acquisition models, he describes the high cognitive effort to control a new skill during an initial phase of learning and skill acquisition, the shift to a smoother performance that no longer requires the previously large amount of effortful attention to perform the skill at a satisfactory level and finally a shift towards automated performance. He emphasizes however that for expert performance, learners must continuously challenge themselves through deliberate practice to continue to improve their skill and avoid staying in automaticity and arrested development (Ericsson 2000: 195–198). Krampe & Ericsson (1996) again highlight the importance of deliberate practice for the maintenance of expert performance. They compared the cognitive-motor skills of young and old amateur and expert pianists. While older expert and amateur pianist displayed a normal age-related decline of general processing speed, their performance on music-related tasks did not decline as much as the performance of old amateurs. Most importantly, the amount of deliberate practice older experts engaged in during the last 10 years of their life prediced how much of the pianistic skill could be maintained. The authors suggested that through deliberate practice, expert pianists acquire and maintain certain mechanisms that enable expert performance in the musical domain. If maintained, these mechanisms make the experts less vulnerable to cognitive decline in the processing skills related to their area of expertise. The study paints out explicitely that deliberate practice is not only important in the acquisition phase of a skill but constitutes what makes expert performance so special and that maintaining expert performance is not an automatic byproduct of having a acquired a skill on expert level once in younger years (Krampe & Ericsson 1996: 348–357).

As described in the previous section, interpreters need to juggle different overlapping cognitive tasks: the overlap between speech comprehension, transfer and production. Chabasse (2009) notes that this overlap requires the interpreter to apply strategies to make the process of interpreting easier and to distribute available cognitive resources economically (Chabasse 2009: 71). Gile (2009) refers to the these strategies as coping tactics and describes them as solutions implemented by the interpreter to deal with the various problems that can arise during SI due to for example cognitive overload (Gile 2009:191-200). Riccardi (2005) characterizes these strategies as knowledge-based strategies since applying them results from conscious and analytical processes when an action during interpreting needs to be planned on the spot as no automatic response is available or too much cognitive overload is already present. She distinguishes these knowledge-based strategies from skill-based strategies, which are defined as stored automatic response patterns that are applied unconsciously such as recurrent and stereotypical parts of speeches (e.g. welcoming, greetings, thanks) (Riccardi 2005:760-762). Corresponding to this view, Andres (2013) also highlights that skills play an important role in SI expertise research, with skills being processes that aid in completing an action and are acquired through training but are then mostly automated (Andres 2013: 239). Hence, both strategies and skill play an important role in SI competence and expertise acquisition and must be trained through deliberate practice (Andres 2013: 240). This notion of deliberate practice of skill in SI corresponds to the aforementioned account of deliberate practice in other domains by Krampe & Ericsson (1996) to maintain expert performance. However, it seems like strategies play a bigger role in conscious decision-making during the SI process when dealing with cognitive load and in restructuring knowledge.

Within the presentation of his Cognitive Load Model, Seeber (2011) addresses different strategies by focussing on cognitive load during interpreting. Similar to Gile’s Effort Model, Seeber’s Cognitive Load Model tries to depict the cognitive demands during interpreting. However, Seeber’s model attempts to quantify cognitive load and enables a more in depth analysis of local cognitive load caused by different interpreting strategies. Seeber lists waiting, stalling, chunking and anticipation as strategies during simultaneous interpreting. During waiting, the interpreter stops producing the target language and waits for more information in the source language whilst storing this information in working memory where it needs to remain in an activated state or be reactivated until further encoding and production. While this strategy eases cognitive load for a certain time, it also creates a spillover effect. This means that cognitive load will eventually increase even more when the speaker produces more and more output. The load is therefore rather delayed than solved. Stalling is similar to waiting in that regard as it also gives the interpreter time to accumulate more information and context. However, interpreters are not silent during stalling as they are during waiting but fills the gap between meaningful utterances with repetitions for example. Stalling makes processing more complex and increases the interpreter's lag. During chunking the interpreters divide a sentence into smaller units that should roughly correspond to the prosodic pattern of the respective language. Thereby, information can be encoded and embedded immediately without waiting for the end of the sentence. As a downside, chunked sentences can sometimes seem convoluted. Lastly, anticipation means that the interpreter tries to predict an utterance before it has been finished by the speaker. The big advantages of anticipation are that the cognitive resources do not increase drastically compared to the other strategies and that the spillover effect is avoided. Only inference processing during anticipation costs some cognitive load (Seeber 2011: 189–195). Seeber (2011) argues that all strategies but anticipation cause a spillover effect and that cognitive load is dependent on how much restructuring interpreters engage in (Seeber 2011: 197).

For the comprehension process, Chabasse (2009) lists strategies such as preparation in terms of content as well as terminology to make inferencing through context as efficient as possible. Knowing the context and being able to infer meaning is crucial for another strategy: anticipation. If the interpreter is well prepared, they can either rely on grammatical indicators that contain meaning or infer meaning through knowing the context and consequently anticipate what the speaker is going to say next (Chabasse 2009: 72). Another strategy is making use of mental frames or scenes. These concepts come from Fillmore's (1976) frame semantics approach. Chabasse (2009) summarizes that frames are the abstract mental representation of a situation and scenes are concrete mental images to that situation (Chabasse 2009: 67). Activating such as frame and scene can ease the cognitive load during interpreting and free resources that they can invest otherwise (Chabasse 2009: 74). Another crucial strategy for comprehension is disambiguation by linking the meaning of an utterance to context. Ambiguity is resolved by suppressing more and more meaning representations that do not fit the conversational context (Chabasse 2009: 88). Chabasse (2009) does not explain the cognitive nature of this suppression process, however, this might be a part of inhibitory executive functions. Chabasse (2009) furthermore describes that the interpreter needs to decide on the meaning of the utterance before receiving the full set of elements needed for disambiguation while still keeping other possibilities in mind where a sentence meaning might be headed. The interpreter cannot wait too long to solve ambiguity because they would risk stunting speech production and forgetting parts of the utterance. The interpreters still need to be able to correct themselves if the decision they took on disambiguation turns out to be false. Chabasse (2009) emphasizes that this process requires immense cognitive flexibility, another core executive function often mentioned in connection with shifting and switching between tasks (see Section 4.2). Furthermore, Chabasse (2009) mentions that interpreters need to keep their closely linked language systems of L1 and L2 from interfering with one another by suppressing the target language during comprehension and suppressing the source language during speech production all while keeping information in working memory. She highlights that these processes have nothing to do with linguistic competence but with cognitive control or attentional control (Chabasse 2009: 89–90). As discussed in Chapter 5. and Section 5.1, it is likely that both languages are active at the same time during bilingual language production and that access to the mental lexicon happens non-selectively. Although it seems unlikely that a whole language system is suppressed, it is still debated whether a local suppression of non-target language representation occurs or whether it is an interplay of increasing and decreasing activation (or inhibition) of language representations (see Section 5.1).

Similar to what Chabasse (2009) describes as disambiguation, Gile (2009) lists reconstructing a segment with the help of the context as an important comprehension strategy. In the case that the interpreter has not completely understood a name or technical term, they can try to reconstruct the information by using their extra-linguistic knowledge (knowledge of the subject or situation) (Gile 2009: 201). Gile (2009) furthermore explains that when interpreters are confronted with memory overload, they can apply the strategy of segmentation to unload the short-term memory. This entails reformulating speech segments earlier than usual, even when the full picture of the sentence is not yet clear (Gile 2009: 205). Moreover, if the interpreter cannot think of a term on the spot, they can use explaining or paraphrasing as a sort of expansion instead of a one-to-one translation (Gile 2009: 207). When a piece of information seems to have negligible value and other crucial information requires the interpreter's attention, they can use the strategy of omission or compressing the content. This does not mean that information is lost, it can surface again at another point if the interpreters choose. However, Gile (2009) emphasizes that this strategy should not be chosen lightly since it can be unethical to omit information (Gile 2009: 210).

Regarding speech production, Chabasse (2009) mentions the four steps of conceptualization, formulation, articulation and monitoring. She highlights monitoring as an important strategy in speech production during SI and, although she does not explicitly mention the mechanisms behind monitoring, she refers to it as the control of the interpreter's output (Chabasse 2009: 57). Although this does correspond to the monitoring mechanism in Levelt's (1989) model of speech production, it is not the same kind of monitoring that is discussed in executive function research, as already discussed in this section.

Lastly, Chabasse (2009) mentions that scenes and frames help to structure content coherently. If this coherence is missing during a speech, the interpreter can use the strategy of inferring by accessing their world knowledge or by recalling so called chunks from memory (Chabasse 2009: 67).

Riccardi (2005) groups the aforementioned strategies into categories: while she assigns anticipation, segmentation, information selection and stalling as comprehension strategies, compression, expansion, restructuring and pauses or intonations are categorized as production strategies. Furthermore, she sees décalage and monitoring as overall strategies and omission for example as an emergency strategy (Riccardi 2005: 765).

The aforementioned strategies that are a part of interpreting competence acquisition highlight especially the cognitive part of interpreting competence. Integrated into other aspects of interpreting competence, they are represented in different models of interpreting competence that are discussed in the following.

Kalina (2000) explained that translation studies[3] have a longer history than interpreting studies which is probably why interpreting was initially seen as just an oral form of translation (Kalina 2000: 16). Interpreting and translating do share the aspect of reformulating from a source language into a target language as translation proper or interlingual translation as Jakobson (1959) calls it, although a message can also be translated or interpreted intralinguistically (reformulated within one language) or intersemiotically (interpretation of verbal signs) (Jakobson 1959: 233). Therefore, translation and interpreting also share various competences that are represented in different translation competence models such as in older models as by PACTE group et al. (2003) and Göpferich (2009) or in a new and much more comprehensive framework of the European Masters in Translation competence framework (see European Commission (2022)) or the EFFORT project (EFFORT Project 2023). Surely, there are important competences in all these models and frameworks that are valid for both translators as well as interpreters. Interpreters and translators do both need for example bilingual sub-competence, instrumental sub-competence and extra-linguistic sub-competence (PACTE group et al. 2003: 58–59) or, as Göpferich (2009) called it, communicative competence, tools and research competence and strategic competence (Göpferich 2009: 20). However, neither of these competence models appropriately covers the idiosyncrasies of the interpreting process mentioned in the beginning of Chapter 6.

In contrast to models of translation such as by the PACTE group et al. (2003) or Göpferich (2009), models of interpreting such as by Gile (2009), Moser (1978) or Seeber (2011, 2017) have rather focused on processes than on competence. However, different lists of interpreting competences have emerged over time (Tiselius & Hild 2017: 428). For example, based on previous literature, López Gómez et al. (2007) listed several “aptitudes” divided into “cognitive” and “personality” categories. The cognitive category includes both linguistic and cultural aspects and comprises efficient input segmentation, attentional division, use of language-pair specific strategies, mastering the predictable properties of language, being able to change translation strategies, possessing verbal fluency as well as processing speed,

---

[3] “Translation studies” is usually used as a cover term for both translation and interpreting studies, however, in this context it is used to describe translation research

good long- and short-term working memory and power of concentration. The personality aptitudes include both cultural as well as so called “academic” aspects and include being able to adapt without delay to different speakers, situation and subjects, having a pleasant voice and public-speaking skills, displaying stress resistance and self-control, being able to work in a team, keeping a professional distance, having a desire to be well informed, being diplomatic and having a good self-concept (López Gómez et al. 2007: 77). A much more specific division was undertaken by Kalina (2006), who separated interpreting competences into pre-process competences, peri-process competences, in-process competences as well as post-process competences (Kalina 2006: 257). This model was taken up by Albl-Mikasa (2012) who used interpreter’s real live experiences to construct a model of interpreting competence. Pre-process competences include mastering languages that are used during interpreting, terminology management, informed-semi knowledge about the topic they are interpreting and a streamlined preparation that itself entails the sub-competences of fast orientation, selective approach, and individual memory support drafting. Peri-process competences entail being able to work in teams and cooperate with colleagues and clients, keeping the balance between actively promoting communication and operating in the background, cultivating an instinct for what is at stake and what lies at the core of an interaction, being realistic about one’s own performance and abilities as well as being able to build up resilience against the pressure of the task and the working field. Embedded in the peri-process skills are the actual in-process skills that are required during the process of interpreting. Here is where many of the aforementioned strategies to manage the increased processing demands and alleviate cognitive effort are located. In-process skills are divided into comprehension skills, transfer skills and production skills. Comprehension skills encompass the competence to scan, identify and match the content of an utterance as well as contextualize what is said. A special mention is given to dealing with interpreting situations where English is used as a lingua franca and mixed into another language by a speaker. The interpreter needs the skill to estimate how much of the English expressions are in fact common knowledge or need to be articulated otherwise. The transfer competences entail the skill of coordinating the overlapping processes of listening, transfer, speaking and monitoring as well as being able to relieve mental load by applying processes such as chunking (see Chapter 6.). Production competences entail modulating the synchronicity and décalage, making minimal reductions to the speaker's output to fee up mental capacity, balancing high fidelity to the source utterance and target audience needs as well as adapting the performance, presentation and prosody to the occasion and the audience. Lastly, the post-process competences concern terminology work after the interpreting job as well as quality control. Albl-Mikasa (2012) extended Kalina's (2006) competence model by also adding the para-process competences that relate to the status of an interpreter as entrepreneur that has to do with customer relations. The competences therefore include business knowledge, knowledge about customer relations and professional standards, the willingness to engage in life-long learning and lastly the competence to be aware and reflect upon one's own operation pattern, solution finding process and extent of knowledge (Albl-Mikasa 2012: 64–89).

In regard to Kalina's (2006) and Albl-Mikasa's (2012) models, the In-process skills are most relevant in the context of this dissertation, since they are most likely to implicate the executive functions. More specifically and as explained in Sections 5.1 and 5.2, executive functions are required during the overlapping word/language comprehension and production processes.

The two models mentioned above are quite comprehensive in their skills, however, they start out from the point of professional interpreters. What is missing to emphasize the development of competence, and the formation of skill is a model that is interpreter training oriented. A very early distinction between novice and expert in different steps was presented by Hoffman (1997). He distinguished between naïves who do not have any knowledge of the conference interpreting domain, novices who are just beginning the SI studies and have already been tentatively exposed to the domain, initiates who are already beyond the introductory stage of the novice and have begun to gather some experience in SI, apprentices, who are well immersed in the domain between one and 12 years, journeymen, who are able to work as a conference interpreter and finally, experts, who are characterized by effective and efficient performance and have gathered specialised skills and knowledge from sustained experience. This model highlights the stage shifts that occur during the development of interpreting expertise (Hoffman 1997: 199–200). He evidently does not differentiate in terminology between competence and expertise and equals years on the job as an indicator expertise. Moser-Mercer (2008) also criticizes that models developed for experts do not cover how skill acquisition in interpreting through various stages proceeds and that novice performance requires specific factors for it to proceed to expert performance. From a performance psychology perspective, she highlights that altough ability and aptitudes do play an important role in interpreting skill acquisition, aspects such as motivation to learn a new effortful skill, the opportuntity of an appropriate physical work environment to perform the skill and the acquisition of domain-specific knowledge significantly contribute to the formation of SI capacity and ultimately, expert performance. She also promotes the shift from teaching methods relying on routine expertise to so called adaptive expertise. This means to shift from the ability of quickly solve familiar domain-specific problems in routine situations to a more meta-cognitive approach, the ability of adjusting more easily to new and unfamiliar situations, creatively looking for solutions, being willing to deal with uncertainty and continuously improving performance throughout life (Moser-Mercer 2008: 1–9).

A model that precisely considers the finding that SI trainees go through various stages in their skill acquisition process and that focusses on the different subprocesses of SI competence acquisition in training can be found in Kutz (2010), who developed the "Leipziger Kompetenzmodell der Dolmetschdidaktik (LKM)" (Kutz 2010: 12) that specifically responds to the formation of knowledge and skills in interpreting. The competences that are listed for a successful interpreting interaction largely overlap with other models of interpreting competence. What is most interesting in Kutz's (2010) approach are different basic assumptions that underly these competences. Kutz (2010) sees interpreting competence as qualifying a person to interpret on a professional level. Furthermore, Kutz (2010) defines that interpreting competence comprises qualifying a person to perform global control processes as well as different subprocesses. They further explain that these global control processes manage different macro operations such as intentions and goals (Kutz 2010: 198). Although these global processes are not explicitly denominated as executive control functions in this source, the definition that was given corresponds to the general definition of executive functions (see Chapter 4) as well as to the involvement of domain-general or “global” control processes in models of language control such as the ICM (see Section 5.1). From this perspective, Kutz (2010) seems to implicate executive functions as an integral part of interpreting competence. Consequently, it can be assumed that executive control will be affected by the formation of interpreting competence during interpreting training. Kutz (2010) furthermore emphasizes that interpreting competence is formed through three stages that are based on principles of skill learning from the educational sciences: knowledge about action, the ability to act and finally proficiency. In the first stage of competence formation, trainee interpreters gain methodological and theoretical knowledge about the interpreting process. Trainees are supposed to reach a point where they know in theory how interpreting is done, but do not yet have the ability to interpret. This is reached in the second stage through a practical approach and actual interpreting training. The interpreting trainee is then able to perform the process of interpreting but in a very controlled, conscious and effortful manner. Kutz (2010) explains that the terms conscious and controlled reflect slightly different levels of processing. The ability to perform the interpreting process in a conscious manner means that the action is performed while control mechanisms are strongly implicated. The ability to perform the interpreting process in a controlled manner is thought to involve processes that are partly automated but still require higher levels of cognitive control than complete automation of a task. In this second stage, the interpreting trainee is also able to consolidate the theoretical knowledge with practical training (Kutz 2010: 206–207). From a theoretical perspective, the second phase is where executive functions are trained and possibly enhanced since this phase entails the practical application of previously acquired knowledge about action. Using Albl-Mikasa’s (2012) terminology, this phase entails training of in-process competences that require executive functions to coordinate the overlapping processes. How much executive control is required depends on how far competence acquisition has progressed, as becomes clear in the third stage. In the third stage of competence formation, Kutz (2010) describes that interpreting trainees improve their skills to interpret throughout their training process, making them more and more proficient in their actions. More and more interpreting processes are automated and can happen subconsciously. The better the trainees master the competences conveyed and practices through training, the more abilities they build and the more proficient they get. Converting abilities into proficiencies has a beneficial effect on the interpreting performance, since it is more automated, faster, more precise and more ergonomic. Kutz (2010) highlights at this points that abilities only turn into fully automated proficiency with intensive professional training and that proficiencies can be lost if not trained regularly, similar to a muscle (Kutz 2010: 206–210).

Riccardi (2005) highlights the different forms of knowledge that are altered during simultaneous interpreting training. She explains that declarative knowledge is non-automatic knowledge about a task that can be verbalized. Procedural knowledge on the contrary is unconscious knowledge about how to perform a task that is more automatic. Interpreting students must improve their declarative knowledge in different topics such as economy or EU policies. Trainee interpreters must learn new procedural knowledge to automatize different interpreting related processes. The author summarizes that professional interpreters and trainee interpreters differ in how their declarative and procedural knowledge is organized, which results in differences in interpreting performance (Riccardi 2005: 757–758). It could be argued that interpreting trainees only possess declarative knowledge about interpreting at the beginning of training in what Kutz denominated the knowledge about action phase. In the ability to act and proficiency phases, the students transform the declarative knowledge into procedural automatic knowledge.

The above described modelling of interpreter trainee competence acquisition overlaps with models of translation competence acquisition such as that by PACTE (2000) which assumes that translation competence consists mostly of sub-competences. The novice stage of acquisition is seen as a stage where certain sub-competences have been partially acquired but are not yet fully developed and do not yet interact with one another. On their way to expert performance, translators restructure these sub-competences (PACTE 2000: 103).

Furthermore, this modelling of interpreter trainee competence, especially the different phases and the eventuating loss of an ability in the absence of usage, corresponds perfectly to other models of skill acquisition such as by Chein &

Schneider (2012) (see Section 2.1) as well as to recent neuroscientific models of functional and structural changes in the brain resulting from language learning (Pliatsikas 2020) (see Section 3.1):

So far, this dissertation has presented three dynamic models that depict different stages of skill learning in general (Chein & Schneider (2012)), language learning (Pliatsikas (2020)) and interpreting (Kutz (2010)). All models include domain-general executive functions and all models include three similar phases of low skill or initial skill formation, performing the skill by help of controlled execution and finally an automation of the skill (see Table 1). In view of this dissertation that asks how executive functions change during interpreting training, it is therefore helpful to theorize based on these three models how interpreting students might progress in their skill learning.

Table 1: Overview of the different stages of three skill acquisition models

Source	Model	Stage 1	Stage 2	Stage 3
Chein & Schneider (2012)	Skill acquisition model	formation	controlled execution	automation
Pliatsikas (2020)	Dynamic restructuring model	initial exposure	consolidation	peak performance
Kutz (2010)	Leipziger Kompetenzmodell der Dolmetschdidaktik (LKM)	knowledge about action	ability to act	proficiency

While the models by Chein & Schneider (2012) and Kutz (2010) assume the most activation of executive functions in Stage 2, when the skill is performed more regularly but has not yet been automated and requires controlled execution, Pliatsikas' (2020) dynamic restructuring model of L2 acquisition sees the initial exposure stage as the stage that requires most cortical executive functions. He mentions however, that Stage 2, where bilingual proficiency has already increased, still requires executive functions, since a constant inhibition of the non-target language and switching between languages takes place (Pliatsikas 2020: 465). The shift from Pliatsikas Stage 2 to Stage 3, which represents peak performance and automation, and engagement of executive functions of course depends among other factors on how often both languages are used and how much switching between the languages occurs.

Back to the interpreting students in training, they should already possess high language competence in both their L1 and L2. In terms of the models, the students would therefore start out at Stage 2 or between Stage 2 and 3 in terms of how consolidated they are in their language skills at the beginning of their interpreting training. According to the dynamic restructuring model, they would have surpassed the initial benefit in executive function that is provided by the initial exposure to L2 and the associated potential increase in grey-matter volume. However, they should still engage their executive functions to manage multiple language systems. Contrary to the language competences, their interpreting competence before starting training should be in Stage 1. In terms of Kutz's (2010) model, they have knowledge about simultaneous interpreting and may be initially exposed to it in the first weeks of training. The SI students should then shift to Stage 2, ability to act, where they require a drastic amount of executive functions to accommodate the new tax on the linguistic system. Even if they have trained their executive functions through bilingualism or translating between L1 and L2, the stress that the simultaneity and online processing put on the linguistic system should be enough to show an increased engagement of executive functions. Through a controlled execution with high engagement of executive functions, the students increase their SI competence. According to Kutz (2010), extensive professional training is necessary to reach Stage 3 of the LKM model, therefore it cannot be assumed that the interpreting students will have fully automatized their interpreting skills. Although they might be proficient performers at the end of training, they are not yet considered experts according to e.g. Hoffman (1997) or Tiselius & Hild (2017).

Another aspect of interpreting competence acquisition is the discussion around whether there is a presupposition for interpreting and the associated aptitude testing before starting interpreting training. SI was long seen as a skill only some individuals have an aptitude for, resulting in aptitude tests to select promising candidates for SI training (see e.g. Moser-Mercer et al. (2000)). However, as Chabasse (2009) emphasizes, training can compensate initially lacking aptitude and testing aptitudes too early in a learning process may potentially exclude individuals whose aptitude needs longer to develop or might set in later (Chabasse 2009: 55). Chabasse furthermore summarizes that by practicing interpreting repeatedly, the brain of the interpreter reorganizes which makes interpreting easier. Therefore, interpreting is an activity that needs to be trained which mitigates the significance of aptitude tests. She highlights that aptitude tests should not measure what an individual can already do, but what they have the potential to be able to do in the future. Students that apply for interpreting training should not be required to already be able to interpret, since this is a skill that is developed through interpreting training (Chabasse 2009: 124). Still, especially when empirically studying executive function by comparing interpreting trainees with other bilingual populations at the beginning of their respective training, it is sensible to draw a baseline comparison of cognitive skills. This way it is possible to observe whether both groups have similar cognitive prerequisites or whether SI students possess inherent advanced cognitive skills as it was done for example by Babcock et al. (2017).

So far, this section has presented strategies that can be applied to manage the processing demands created by the task of simultaneous interpreting, presented models of interpreting competence in which these strategies are integrated as well as models of interpreting competence acquisition. Especially the restructuring of knowledge during SI competence acquisition has been a central theme. This restructuring of knowledge and the whole process of learning to interpret has also been studied empirically with studies finding effects of emerging interpreting competence on linguistic, cognitive, and interpreting-specific performances:

Older studies such as by Moser-Mercer et al. (2000) found that expert interpreters outperformed novices in a delayed auditory feedback (DAF) task. During DAF, a delay between the moment a speaker produces a speech sequence and the moment the speaker hears the output is created electronically. Results showed that experts made fewer mistakes in delayed auditory feedback and displayed greater verbal fluency than the novice interpreters, probably since the experts needed to pay less attention to monitoring their output as a result of their increased SI competence (Moser-Mercer et al. 2000: 123).

Bajo et al. (2000) compared the performance of professional interpreters, interpreting students, bilinguals and non-interpreting professionals from other fields in different linguistic abilities such as reading time, lexical decision and semantic categorization. They hypothesized that interpreters would outperform all other groups since they have specifically trained their comprehension skills in addition to their bilingual skills. Results showed that interpreters showed significantly faster reading times than bilinguals and non-interpreters, indicating increased information processing efficiency. For the lexical decision task, results showed significant differences in reaction times between groups for non-words with interpreters outperforming all other groups in reaction speed. In terms of semantic categorization, interpreters performed faster in non-typical concept categorization than all other groups. To see if the effects of superior performance in interpreters were due to training, the interpreting students had to perform the tasks twice, once at the beginning and once at the end of an academic year. A control group with non-interpreting students was tested in parallel. Results indicated that reading times reduced significantly from Session 1 to Session 2 for the interpreting students but not for the control group. Similar result patterns were revealed in the semantic categorization task: Interpreting students responded faster for both typical and non-typical items in Session 2, while non-interpreting students did not seem to improve their speed of accessing semantic knowledge. For the lexical decision task there was only a tendency for interpreting students to improve their response speed, the difference between Session 1 and 2 was however only significant in the non-word condition. No difference in reaction times between Sessions 1 and Session 2 surfaced. In another separate experiment, Bajo et al. (2000) also tested the groups on their working memory capacity by administering a digit span task, a phrase span task and a word span task with and without articulatory suppression. Results showed that interpreters had significantly higher memory spans in the digit and phrase span tasks compared to bilinguals and non-interpreting professionals. As indicated by the results, interpreters, students, and non-interpreters performed similarly in the condition without articulatory suppression. However, in the articulatory suppression condition, professional interpreters outperformed the other groups. This indicates that extensively trained interpreters have better working memory capacity and can use their working memory more efficiently. The authors conclude that interpreting training and professional interpreting practice enhances different linguistic and non-linguistic skills that are connected to comprehension. Interpreters must therefore develop a special set of skills (Bajo et al. 2000: 130–140).

Tzou et al. (2012) examined the influence of language proficiency and formal simultaneous interpreting training on interpreting performance and working memory by comparing interpreting students with one year of formal training experience, interpreting students with two years of formal training experience and a non-interpreting bilingual group. For measuring working memory performance, they used a spoken digits span and a reading span task. A simultaneous interpreting task was used to determine interpreting performance, recordings of the task were rated by professional interpreters. Language proficiency was measured with a TOEFL test. The digit span task showed a significant effect of group with second year interpreting students showing significantly higher scores than non-interpreting bilinguals. The reading span also showed a significant effect of group, this time both first- and second-year interpreting students outperformed the bilinguals. Results of the interpreting task showed a significant effect of group with the second-year interpreting students outperforming the first-year interpreting students in overall quality. The first years outperformed the non-interpreting bilingual group. Concerning the role of language proficiency in relation to interpreting and working memory span, results revealed that advanced L2 proficiency did not enhance working memory spans more across the two languages compared to less advanced L2 proficiency. Lastly results showed a positive correlation between L1 and L2 reading span and simultaneous interpreting quality and between L2 digit span and interpreting quality. Taken together the results show that different stages in interpreting competence acquisition are reflected differently in interpreting quality and that working memory span is partly associated with interpreting competence and quality (Tzou et al. 2012: 217–225).

Within the scope of an fMRI study to longitudinally investigate brain changes in simultaneous interpreting trainees, Hervais-Adelman et al. (2015a) also compared simultaneous interpreting performance. Participants were tested at the beginning and at the end of their training and were compared to a non-interpreting control group. Results showed that interpreting trainees outperformed the control group both at the beginning and at the end of training in the interpreting task. The difference between groups at the beginning of testing could be attributed to the fact that they self-selected themselves for the interpreting training according to their talents and interests. However, there were also significant improvements within the interpreting group from beginning of training to the end of training. Response latencies showed that interpreters performed faster in the second session compared to the first session (Hervais-Adelman et al. 2015a: 268). fMRI results are discussed in Section 6.3.

Chmiel (2021) showed how interpreting training and the associated competence acquisition affects word translation speed by testing interpreting students at the beginning and at the end of their training programme in a sentence translation task and compared the performance to professional conference interpreters. The translation task included context (high constraint context, low constraint context, no context) and translation direction as independent variables. They expected translation accuracy and translation speed to increase as a result of training. More specifically, translation into L1 should be faster than translation into L2 at the beginning of training. This direction effect was expected to disappear at the end of interpreting training, and it was expected to not be apparent in professional interpreters, since the interpreting training and the professional practice affect the competence of switching between the languages. Results showed that interpreting students differed significantly in their translation accuracy before and after training and that trainees generally performed better when the words were embedded in sentence context both before and after training. Accuracy was modulated by context in the L2-L1 direction but not in the L1-L2 direction. Regarding translation speed, results showed that interpreting trainees were faster after training than before training and were faster in the L2-L1 direction than in the L1-L2 direction in general. Professional interpreters were compared to the interpreting students at the end of their training. Professional interpreters showed maximum accuracy. Due to this ceiling effect, it is not possible to draw conclusions regarding accuracy, although the professionals did outperform the students. There was also no main effect of experience on translation speed when comparing students at the end of their training and professional interpreters. This might also represent a ceiling effect since the cross-lexical links might have already become reinforced at advanced pre-professional training so that professional training cannot give a further boost. Professional interpreters still showed an advantage in the L2-L1 direction against expectations. Interestingly, the low constraint context cost decreased with experience. Chmiel (2021) interpreted this observation in favour of a more efficient inhibition mechanism to select the correct translation equivalent over time. Overall, this study gives a good example of how to differentiate between the effect of training and the effect of professional experience in interpreting and also tentatively discusses the involvement of the executive function of inhibition (Chmiel 2021: 25–38).

Yu & Dong (2022) investigated how interpreting competence develops by longitudinally examining how interpreting competence is correlated with language competence and working memory capacity in consecutive interpreting students. They orient their theory to the dynamic systems theory that assumes that complex systems such as human development is a non-linear process that has no definite end. More traditional developmental theories assume that development proceeds predictably and linear and has a clear beginning and end. They postulate that with increasing interpreting competence, trainees can better coordinate various tasks. Therefore, they hypothesize that students have increased cognitive abilities after training that correlate with interpreting performance. Students performed consecutive interpreting tasks, five language competence evaluation tasks and 4 working memory assessment tasks before and after 10 months in interpreting training. Although the authors speak of cognitive abilities, most tasks here are language specific with the only executive function specific tasks being the working memory assessment tasks. Results showed overall that with more interpreting training, more language and memory skills were correlated with interpreting performance. They summarized that within the first year of interpreting training, students seemed to be able to transform their explicit, declarative knowledge into implicit procedural knowledge. This study, however, did not examine how other cognitive control functions interact and contribute to interpreting performance (Yu & Dong 2022: 269–280).

These studies have shown that the acquisition of interpreting competence gradually affects different language specific abilities that are trained during interpreting training as well as interpreting specific abilities. The times lines that were examined were quite different. One year of training (Tzou et al. (2012)) or one academic year of training (see Bajo et al. (2000)) already leads to some changes and Chmiel (2021) investigated changes after a two-year training programme. It has become clear however that changes already occur during interpreting training, as is reflected in relevant models of general skill acquisition and interpreting skill acquisition (see e.g. Kutz (2010)), and not only after several years of professional experience. The remaining question that I will address in the methodological part of my thesis is at what point in training does competence affect certain abilities and is an increase of abilities normalized within the time frame of interpreting training.

One issue that studies investigating interpreting competence acquisition have in common is the assessment of competence in translation and interpreting process studies. This is an issue related disciplines such as bilingualism research also face, as laid out in Chapter 3. Measuring or categorizing bilingualism encounters challenges such as heterogenicity in methods or approaching bilingualism not as a continuum but as something static which makes comparison across studies difficult. In the field of translation and interpreting studies, García (2014) points out that most studies use years of professional experience as the key indicator to measure expertise or competence. He criticizes, however, that the years of experience do not actually say anything about how often an interpreter is on the job or practices, nor does it say anything about the quality of their interpretation. He also emphasizes that large variation in professional experience can make a measure unreliable. He proposes that other measures should be considered when gauging interpreting competence and expertise, such as hours of professional practice, work settings, types of training or interpreting quality ratings (García 2014: 228–229). Schaeffer et al. (2020) also highlight that "[…] measures of translation and interpreting competence in relevant experiments are either absent or based on non-validated ad hoc instruments." (Schaeffer et al. 2020: 91). Schaeffer et al. (2020) therefore presented the Translation and Interpreting Competence Questionnaire (TICQ), a validated questionnaire that makes it possible to collect quantitative and qualitative data on different aspects of translation and interpreting competence as well as bilingual proficiency such as age of acquisition, speaking habits and self-assessed proficiency. This questionnaire was also used in the experimental setup of this dissertation (for a detailed description see Section 8.1.3).

Finally, it is vital to mention that the presented models of competence and competence acquisition emphasize the cognitive side of simultaneous interpreting competence since this dissertation is first and foremost interested in the domain-general cognitive processes that are affected by SI. However, there is of course more to competence and expertise than the cognitive side of things. This is emphasized by Beaton-Thome (2018), who complements the cognitive aspects of SI competence by highlighting the situatedness of SI and the importance of situated learning, situated expertise and community of practice. Beaton-Thome (2018) recognizes it is sensible to see one part of SI competence and expertise situated in the mastery of certain cognitive skills such as multitasking, monitoring, chunking and adjusting the décalage, since this is what can be tested and verified in cognitive interpreting studies. She argues however, that viewing how much time an interpreter has spent on a task should not be equalled to cognitive skill, since there are other parts of expertise that play a crucial role. She emphasized the concept of situated learning, which means learning within the context of a group and thereby participating in a community of practice. She argues that this community-based approach to learning should be intersected with and integrated into cognitive aspects of SI expertise. To this end, she proposes different strategies to implement this already during interpreting training and competence acquisition. One option could be to instate a mandatory number of days of work experience as a conference interpreter that the student trainees need to experience under the guidance of an experienced interpreter. Also, she proposes student-tutor booth constellations. Most relevantly in the context of this dissertation, she also suggests that this situated learning within a community of practices should occur from the beginning of training throughout all competence acquisition phases and not only at the end of training, which could provide a smoother transition from novice to proficient user to expert, or in other terms, a smoother transition from competence acquisition to expert performance and expertise (Beaton-Thome 2018: 148–160).

Section 6.1 presented cognitive models of the SI process and highlighted the challenges interpreters face arising from overlapping cognitive processes and the involvement of executive functions. This current section has highlighted the role of SI strategies in competence acquisition to deal with the challenges of the SI. Furthermore, the notion of competence was differentiated from the notion of expertise. Moreover, this section has highlighted which competences are important to the interpreting process, where executive functions are implicated in these competences and how interpreting competences are acquired. Thereby, the overlap of interpreting competence formation with general skill acquisition models and recent findings from neuroscience concerning the consolidation of competence has been specified. Studies emphasizing that simultaneous interpreting is a skill that needs to be acquired and that the formation of competence and training have effects on different cognitive skills such as working memory, semantic categorization or lexical decision were presented (see e.g. Tzou et al. (2012); Yu & Dong (2022); Bajo et al. (2000)). This again highlights the impressionable nature of the brain. However, the studies presented above mostly focused on how interpreting training fosters an increase in interpreting competence in general or how interpreting training improves the above mentioned linguistic functions in a behavioural manner. The next section therefore continues this train of thought and examines from a neuroscientific perspective how the brains of interpreters change due to the demanding task of interpreting, presenting evidence for the involvement of executive function in the SI process.

6.3 The neuroplasticity of simultaneously interpreting brains

Chapter 6 so far has pointed out that during simultaneous interpreting, various cognitive processes overlap, and different models of simultaneous interpreting have tried to model the cognitive effort and control that needs to be exerted to achieve a simultaneous translation. Since neuroscientific methods have also become more available in cognitive translation and interpreting research, various studies and theoretical frameworks have set out to identify the demanding bilingual processing at work during simultaneous interpreting, exploring the cognitive basis of simultaneous interpreting and investigating how brain structure and functionality are already subject to change as a result of simultaneous interpreting. As Chapter 3 has shown, managing more than one language has shown to elicit structural and functional changes in the brain. The following studies provide a chronological overview of studies using both behavioural as well as neuroscientific measures to show how the brain changes as a result of simultaneous interpreting as a form of extreme language control:

In one of the earliest studies that revolved around the cognitive processes of simultaneous interpreting, Fabbro & Darò (1995) hypothesized that because of the complexity and the simultaneity of the processes of simultaneous interpreting, interpreters specifically need to develop the ability to focus their attention on the verbal input and only direct a reduced focus to the output while displaying increased verbal fluency to keep up with the respective speaker and still have time for potential self-corrections. They tested whether simultaneous interpreters possess increased verbal fluency by conducting a study during which they compared monolinguals and interpreters in a verbal fluency task under normal and delayed auditory feedback conditions. Delayed auditory feedback has been found to cause speech disruptions, however, these speech disruptions are usually less frequent in speakers with high verbal fluency. They found that monolinguals performed significantly better in the normal auditory feedback condition than in the delayed auditory feedback condition while there was no significant difference between the number of words produced during normal auditory feedback and delayed auditory feedback in the interpreting group neither for L1 nor for L2. The authors conclude that simultaneous interpreters have learned to focus less on their own verbal output and direct their attention rather on the input message that needs to be translated and are therefore less susceptible to speech disruptions (Fabbro & Darò 1995: 315–316). Although it is not specifically mentioned in their study, the fact that interpreters were better at ignoring delayed auditory feedback fits with the bilingual advantage research that found increased abilities of bilinguals to suppress interferences.

Rinne et al. (2000) used PET to explore brain activation in professional simultaneous interpreters during tasks of interpreting, either translating from their L1 into their L2 or vice versa and contrasted these activation patterns with shadowing. During interpreting into L1, results showed increased activation on the left frontal lobe. Interpreting into the L2 was characterized by even more pronounced left frontal activation as well as left inferior temporal activation (Rinne et al. 2000: 86–87). These findings highlight the importance of the left frontal regions that are associated with executive functions, especially during interpreting into the non-native language which can pose more difficulty when L2 and L1 proficiency are unbalanced.

Within the scope of an EEG study, Elmer et al. (2010) investigated how long-term language training such as in simultaneous interpreting training influences neural adaption in single word processing. The goal was to distinguish the influence of proficiency and age of acquisition from the influence of long-term interpreting training. They used the N400 ERP component that is sensitive to lexical-semantic processing and semantic anomalies. The N400 amplitude is smaller for words that are semantically related than for semantically unrelated words. Elmer et al. recruited professional SI interpreters that had trained and graduated in a university interpreting programme and exclusively interpreted from their L2 into their L1 as well as non-interpreting controls matched for age, gender, L2 proficiency and AoA. All participants performed semantic decision tasks where they judged whether noun-pairs were semantically congruent or incongruent within and across two languages. The results did not show electrophysiological differences between the two groups in the language direction L2 to L1 which SIs specifically trained for. However, results showed more negative N400 responses in SIs in incongruent trials within L1 as well as L2 but also while they performed the task from L1 to L2 which they did not specifically train for. The increased N400 amplitude in SI might indicate that SI training has altered the SIs' sensitivity to semantic processing within and across both languages. Participants within the SI group differed in years of work experience. More negative N400 amplitudes were revealed in the SI group with less work experience compared to participants with more work experience. The results also suggest a more pronounced interconnection of semantic representation in both languages which causes more lexico-semantic neighbours to be coactivated (Elmer et al. 2010: 148–154).

Ahrens et al. (2010) compared simultaneous interpreting to free speech based on the premise that brain activation patterns of perception/comprehension and speech production differ in the two tasks, since simultaneous interpreting required parallel processing of two languages and is more than just hearing and speaking. The left superior temporal sulcus (STS) was identified as the most active area during simultaneous interpreting. The authors highlighted that the left STS is not only involved in extracting speech and sound that is perceived but also in inhibiting the perception of the interpreter's own voice (Ahrens et al. 2010: 244–245). This speaks for an involvement of a domain-general interference suppression mechanism that is active during simultaneous interpreting.

Elmer et al. (2014) argue that although it is useful to study functional and structural neuroplasticity, executive function and language control in bilinguals compared to monolinguals, studying simultaneous interpreters may deliver important insights into how brain regions related to cognitive control and linguistic functions change in response to training in younger adults. In an MRI study, Elmer et al. (2014) compared grey matter volume in professional SIs to multilingual controls. They found reduced grey matter volume in SIs in the left middle-anterior cingulate gyrus that has been associated with conflict monitoring and cognitive control, bilateral middle-anterior insula and left pars opercularis which have been associated with overt speech production, left inferior parietal lobe that has been associated with attention and working memory and bilateral pars triangularis which has been associated with syntactic processing. Although there were no group differences in grey matter volume in the caudate nuclei, a negative relationship between grey matter volume in bilateral caudate nucleus and the number of SI training hours was found, which suggests that there is a relationship between the amount of SI training and certain grey matter changes. Elmer et al. (2014) indicate that studies in grey matter volumes that compared bilinguals to monolinguals usually found increased grey matter volume for bilinguals. However, they refer to the process of cortical pruning (see Chapter 2) where ineffective neuron connections are removed to foster functional specialization. They argue that excessive training of language control such as during interpreting training in younger adults may lead to supplementary pruning in adulthood, further fostering functional specialization (Elmer et al. 2014: 181–186). Contrary to Elmer et al. (2014), Becker et al. (2016) found increased grey matter volume in the left frontal pole for interpreters when they compared simultaneous interpreters to a multilingual control group. They furthermore found that the left frontal pole was stronger connected to the left inferior frontal gyrus and the middle temporal gyrus in the interpreting group compared to the control group. The authors summarize that the left frontal pole is implicated in monitoring and goal integration in working memory tasks as well as in attention shifting and reallocating executive resources. The left inferior frontal gyrus has been associated with both task switching as well as language switching (Becker et al. 2016: 256–258). More research is necessary to establish under which specific conditions interpreters sometimes display increased grey matter volume and sometimes reduced volume.

In terms of brain connectivity, Elmer & Kühnis (2016) used the EEG to compare the functional connectivity between the left auditory-related cortex and the Broca’s region, two main regions involved in hearing and speaking in simultaneous interpreters and a multilingual control group. The authors predicted that since simultaneous interpreters are well practiced in overlapping hearing and speaking processes, they will pre-activate the connection necessary to anticipate the articulatory code of a target sentence more strongly. Results indeed showed that simultaneous interpreters engaged the connection between auditory related cortex and the Broca’s area earlier and stronger. They also found specific training related changes, with increased functional connectivity correlating to the number of hours the interpreters trained as well as the age at which the interpreters started training (Elmer & Kühnis 2016: 5–6).

Hervais-Adelman et al. (2015b) investigated the neural basis of interpreting and its overlap with domain-general executive functions. They scanned interpreting students before they started their first semester of interpreting training and multilingual controls with an fMRI scanner and compared their brain activity during three different conditions: passive listening (PL), shadowing (SH) and simultaneous interpreting (SI). During passive listening, participants only listen to a recording while during shadowing, participants must simultaneously repeat what is said in the recording. Finally, during SI, participants must interpret what is said in the recording into another language. If the brain region activity engaged in PL and SH are subtracted from the brain region activity of SI, brain regions that are specific to SI should be revealed. Results demonstrated that in addition to all brain regions engaged in SH and PL, SI activated brain regions that manage bilingual input streams in real time. Those regions covered the left anterior SMA and pre-SMA, left anterior insula, left premotor cortex, caudate nuclei and the dorsal ACC as well as the basal ganglia, left pars triangularis and pars orbitalis of the inferior frontal gyrus. The network responsible for simultaneity that was present in SH was activated more expansively during SI, including the superior temporal gyri, the medial prefrontal and medial orbitofrontal cortices as well as the putamen and superior aspect of the cerebellum on both sides. Those regions have been associated with selection and initiation of action sets, error monitoring, multilingual language control, speech preparatory loop, language switching and multitasking. Especially the basal ganglia have been associated with attention, learning, memory and executive functions and have been found to be involved with bilingual language control. Hervais-Adelman et al. (2015b) assume that especially the caudate nucleus is active during SI in monitoring and controlling which lexico-semantic system is needed in the given situation. Their results provide further evidence of how greatly the neural bases of language control and executive functions overlap (Hervais-Adelman et al. 2015b: 4731–4736).

Hervais-Adelman et al. (2015a) used fMRI to specifically examine how the brains of trainee interpreters change functionally during their training to become professional interpreters by looking at cerebral responses. They also used multilinguals as a control group. They hypothesize that increasing interpreting expertise during training progressively automizes language control which shares neural correlates with some aspects of domain-general executive functions. Participants were scanned at the beginning and at the end of a 15 month training period and performed in three conditions: Listening to sentences, shadowing sentences and interpreting sentences. They discovered training-related changes in interpreting performance with interpreting trainees showing a significant improvement in the interpreting condition but no observable changes in the controls. In a univariate analysis they discovered that the response of the right caudate nucleus decreased significantly after the training period in the trainee interpreter group. The caudate nucleus is viewed as a crucial structure in both language and executive control. This result indicates two things. Firstly, that interpreting training induces functional neuroplasticity. Secondly, it suggests that increased competence in interpreting leads to more automaticity in performance which, consequently, requires the use of less neural cerebral resources. Hervais-Adelman et al. (2015a) point out that other studies on experienced interpreters have shown no involvement of the caudate nucleus during simultaneous interpreting, which indicates that the caudate nucleus is engaged less while experience increases. They also discuss the right side lateralisation of the active caudate nucleus in this study, pointing out that previous studies have shown bilateral caudate involvement with a rightward lateralization, although basal ganglia in bilingual language control have a tendency to be left lateralized (Hervais-Adelman et al. 2015a: 268–272).

As already mentioned in Section 3.2, Hervais-Adelman et al. (2017) compared cortical thickness before and after simultaneous interpreting training in interpreting trainees and multilingual controls. Next to increases in cortical thickness in regions of phonetic processing, formulation of propositional speech and working memory, they also found increased cortical thickness in regions of domain-general executive control and attention, namely the right parietal lobule, for the interpreters but not in the control group (Hervais-Adelman et al. 2017: 215–216).

Based on the converging neuroscientific findings regarding the neural basis of simultaneous interpreting, Hervais-Adelman & Babcock (2020) proposed a neurocognitive model of simultaneous interpreting. It takes certain aspects from the model of bilingual language control put forth by Green & Abutalebi (2013) in the context of the adaptive control hypothesis. The model contains distinct control pathways for global task demands of simultaneous interpreting (red) and pathways that are specific to controlling resources for managing multiple languages. According to the author's rationale, the model implies that language depends on domain-general executive functions to control behaviour and that managing multiple languages draws on these functions depending on the context of the language situation and the language status (Hervais-Adelman & Babcock 2020: 747–749).

This section has presented evidence for activity-dependent plasticity of the brain in reaction to simultaneous interpreting. It has become clear that the overlapping cognitive processes at work during simultaneous interpreting affect and modulate brain structure and functionality of both language specific as well as executive functions regions.

This brings me to the last missing puzzle piece of this theoretical part. This dissertation has established a connection between language processing and executive functions (see Section 5.3). Following the conclusions of some researchers of the bilingual advantage hypothesis that frequent language switching leads to an enhancement of executive functions (e.g. Verreyt et al. (2016)), and considering the findings from interpreting competence acquisition that interpreting is an acquired skill that involves executive functions, and taking the evidence from experience-dependent plasticity in interpreters into account that show structural changes in both linguistic and executive function regions, leads to the assumption that simultaneous interpreters should display advantages in executive functions at least during the course of their training. The next section therefore discusses evidence from studies that investigate whether interpreters show an advantage in executive functions by presenting studies that examine both professional interpreters as well as interpreting trainees.

6.4 The impact of simultaneous interpreting on executive functions

As Section 6.3 has shown, there is structural and functional evidence that executive function brain areas seem to be implicated in the simultaneous interpreting process and are therefore also subject to structural and functional changes induced by the enhanced control demands of managing multiple languages. These enhanced processing demands, which also lie at the core of the bilingual advantage hypothesis (see Section 5.4) as well as the knowledge of overlapping brain areas of language control and executive functions and the ability of the brain to reshape itself in reaction to external experiences such as bilingualism, are the basis of the "interpreter advantage hypothesis" (García 2014: 221). It postulates that […] the development of expert interpreting skills may further enhance specific linguistic and executive functions in bilinguals" (García 2014: 221).

The interpreter advantage hypothesis has been tested both behaviourally as well as by using neuroscientific methods, although only few studies have made use of the latter methods. Most frequently, professional or trainee simultaneous interpreters were compared to a multilingual control group or to translators. The research landscape of the interpreter advantage in executive functions is presented chronologically in the following, by listing studies that compared the performance of either interpreting students or professionals to groups of various other bilingual modalities in executive function tasks such as translators, language teachers or bilinguals without special training:

The studies on an interpreter advantage in executive functions revolve around working memory and the question whether professional interpreters have better WM capacity than novice controls or bilinguals without special training.

As already laid out in detail in the previous section, Bajo et al. (2000) compared working memory capacity in professional interpreters to interpreting students and non-interpreting professionals by testing them on a digit span task, a phrase span task and word span task with and without articulatory suppression. Professional interpreters outperformed the other groups in the digit as well as the phase span task. Regarding the word span task, all groups performed similarly when no articulatory suppression was required. However, during the articulatory suppression condition, interpreters recalled significantly more words than subjects in the other two groups (Bajo et al. 2000: 183-140).

Köpke & Nespoulous (2006) compared performance in different working memory tasks in professional interpreters, interpreting students, bilingual controls and monolingual university student controls. They used serial span tasks that only measure short-term retention and only tap into memory storage, more complex recall tasks with articulatory suppression that involve both storage and processing and tap into executive control, listening span tasks, and finally Stroop tests that tap into attention and storage. The serial memory span tasks and the Stroop task did not yield significant group differences, aside from a bilingual Stroop task in French. However, between-group effects were found in the more complex recall tasks with articulatory suppression and the listening span task where novice interpreters performed better than professional interpreters and the professionals performed better than the two control populations. The results showed that interpreting does not seem to lead to simple enhanced short term memory or selective attention but to enhanced cognitive skills that concern the Central Executive. Since interpreters performed better in the articulatory suppression condition than controls, it could be assumed that simultaneous interpreting promotes resistance to phonological interference. It was nonetheless surprising that novice interpreters outperformed expert interpreters which might indicate that differential processes take place in novices and experts. Novices might receive a boost in working memory capacity and suppression skills, while expert interpreters develop other processes (Köpke & Nespoulous 2006: 6–16). This explanation was also suggested by Köpke & Signorelli (2012) who argued that novices tend to have better WM skills since they frequently struggle with cognitive overload and need more executive control than professionals, who have developed more automatic strategies that do not rely on WM as much or not at all (Köpke & Signorelli 2012: 189). Christoffels et al. (2006) examined measures of WM together with basic language processing components of lexical retrieval. They tested experienced professional interpreters and bilingual university students on reading span, speaking span and a word span task as well as picture naming and word translating. Results showed that interpreters were significantly faster in both picture naming and word translation. However, in picture naming this effect only surfaced in English. In Dutch picture naming, students did not differ significantly from interpreters. Furthermore, interpreters outperformed students on all three memory tasks. The results therefore showed distinct differences between the two groups in all the subskills that were assumed to underly the interpreting process. However, since the language proficiency of the two groups were not on the same level, the experiment was replicated with interpreters and language teachers. The results showed that teachers performed similarly to the interpreters on the lexical retrieval tasks, however, the interpreters outperformed the teachers on all three tasks, regardless of the language tested. The biggest takeaway from this study is that interpreters’ WM and short-term memory (STM) capacity was not influenced by L1 or L2 testing, indicating that they have developed a more efficient way of language processing especially in the L2 and that WM is an important subskill of interpreting (Christoffels et al. 2006: 327–339).

These findings that interpreters seem to outperform non-interpreters in WM tasks where more executive functions are involved, solidified the view that interpreting is more than just maintaining information but also involves sustained monitoring, updating, and manipulating of information input and output. Therefore, studies investigating enhanced WM abilities of interpreters started to focus more on tasks that not only tap into WM storage but emphasize maintaining and updating (monitoring) task-relevant information (Morales et al. 2015: 83). One of these tasks is the n-back task that was applied by Morales et al. (2015) when they compared simultaneous interpreters to bilingual controls in the n-back task and an attention network task to test both WM-updating as well as attentional control. Results from the n-back task showed that simultaneous interpreters displayed better monitoring and information updating abilities than bilingual controls. Interestingly, increasing memory load led to decreased accuracy in both groups which led researchers to the conclusion that the storage component of WM functioned similarly in both groups. That interpreters performed better than the control group overall might be due to an enhanced executive component of WM. The difference in performance increased in the second experimental block, where task demands had increased, with interpreters outperforming the controls. This might indicate that the interpreters were able to adapt their performance to the increased task demands more flexibly than the control group. In the attention network task, there were no group differences concerning the executive control network. Similar to the findings by Köpke & Nespoulous (2006), this indicates that an interpreter advantage in executive function might not extend to conflict resolution or that inhibitory processes are not involved or affected by simultaneous interpreting (Morales et al. 2015: 85–89).

The research on working memory and interpreters has shown that, although bilingualism can be an important prerequisite for advantages in executive functions to surface and for the process of interpreting itself, it is the special skills developed during interpreting training that can lead to a change in executive functioning. More evidence for this hypothesis comes from studies that not only investigated working memory span or working memory updating skill but also inhibitory control and task switching skills. The challenge for interpreters is to keep two languages active at the same time while avoiding interference between them. As proposed by Dijkstra & van Heuven (1998) and Green (1998), selection of languages in bilinguals is achieved by inhibiting language representations in the language that is not being used. Furthermore, the executive control function of switching between mental sets (also referred to as flexibility) overlaps with the language switching aspect of interpreting. Therefore, inhibitory processes as well as switching processes have also been investigated in the context of enhanced executive function due to interpreting training.

Yudes et al. (2011) compared professional interpreters to bilinguals and monolinguals. They applied the Wisconsin Card Sorting Test to examine shifting abilities and the Simon task to investigate the inhibition of irrelevant information. Furthermore, they also applied a reading span test which yielded significant differences among the groups with interpreters showing higher memory span compared to monolingual and bilinguals. The results from the WCST showed that interpreters completed the task more efficiently than monolinguals and bilinguals. Interpreters succeeded in updating the task-relevant information more efficiently and were able switch to another solution path when it was necessary compared to the two control groups, suggesting that this advantage in shifting or mental flexibility is related to the interpreting experience. However, the Simon effect in the inhibition task was similar for all experimental groups. No advantage in inhibition for interpreters or bilinguals was detected, which corresponds to the results of Köpke & Nespoulous (2006), raising the question whether inhibition is involved in interpreting at all.

Macnamara & Conway (2014) provided further evidence for an interpreter advantage in executive functions from bimodal simultaneous interpreters. They studied

American Sign Language (ASL) simultaneous interpreting students longitudinally to see if their task switching ability, mental flexibility and two working memory measures changed over the course of training. They found that after their training period, interpreting students outperformed their pre-test scores in task switching, mental flexibility and WM coordination and transformation. However, they did not improve their WM storage and processing abilities. They conclude that an advantage in executive functions comes from an interaction between bilingual management demands and how much experience the interpreters or bilinguals have in accommodating those demands (Macnamara & Conway 2014: 522–524).

Going in a similar direction, Dong & Xie (2014) investigated the impact of language proficiency and interpreting experience on the executive functions of inhibition and shifting in four experimental groups. Two groups consisted of interpreters (undergraduate and graduate) and two groups consisted of Chinese-English bilingual students. The four groups had three different levels of L2 proficiency with graduate interpreters showing the highest level of L2. Results showed that L2 proficiency did not enhance cognitive control in inhibition or shifting in the bilingual students without interpreting experience. This corresponds to Bialystok et al. (2009) who suggest that young adults are cognitively speaking at peak performance which cannot be further enhanced by bilingualism alone. However, the interpreting experience of the interpreters influenced cognitive control in shifting but not inhibition. In contrast to the study by Yudes et al. (2011) where L2 proficiency was controlled, L2 proficiency was manipulated in this study. The results suggest that L2 proficiency is not a main contributing factor for changes in the executive functions of shifting and inhibition, however, interpreting training and experience positively influences the enhancement of shifting but not inhibition (Dong & Xie 2014: 511–516).

Henrard & van Daele (2017) compared translators to interpreters but also compared them to monolinguals in different executive functions tasks that tapped into information processing speed, updating of information in WM, response inhibition, flexibility (shifting) and proactive inhibition. Results showed that interpreters performed better than monolinguals on all tasks and better than translators on all task except on the task assessing flexibility. Translators only performed better than monolinguals in the tasks assessing flexibility and the task that tapped into assessing the resistance of proactive inhibition. This study provided further evidence that an advantage in executive functions is modulated by the cognitive demands of the bilingual situation and that it is likely related to specific aspects in the interpreting work activity such as time pressure and experience in managing the increased processing demands (Henrard & van Daele 2017: 4–9).

Woumans et al. (2015) further investigated the impact of the degree of bilingualism and interpreting experience on the change in the executive function inhibition. They hypothesised that more practice in language control leads to enhanced executive functions. They compared monolinguals, balanced bilinguals, unbalanced bilinguals, and student interpreters, using the Simon task and the Attention Network Task. Furthermore, they used a single and a dual-language version of the semantic verbal fluency task to measure language switching proficiency and relate it to domain-general cognitive control. For bilinguals compared to monolinguals, results produced a smaller Simon effect and faster reaction times in the ANT for bilinguals. Furthermore, bilinguals performed better at orienting in the ANT. In both tasks, interpreters made significantly fewer errors than unbalanced bilinguals but not than balanced bilinguals, providing evidence of an inhibitory advantage of interpreters compared to unbalanced bilinguals. Ultimately, a connection between language switching and cognitive control could only be observed for balanced bilinguals on the Simon task. The authors observed that reaction times and congruency effects were not consistently smaller for bilinguals, which fits right into the argument that different tasks that are thought to tap into the same executive function, actually tap into different categories of that executive function. It has to be noted that the interpreting group were "only" student interpreters, which might explain why there were not more significant differences to balanced bilinguals (Woumans et al. 2015: 1581–1585).

Van der Linden et al. (2018) conducted a similar study where they compared professional interpreters, monolinguals and L2 teachers. They used an advanced Flanker and a Simon task to measure interference suppression and response inhibition, the digit span task for short term memory assessment, the Hebb repetition paradigm and the n-back task for WM updating. Surprisingly, the authors did not find any evidence for an advantage in executive function for interpreters relative to monolinguals or L2 teachers. They also did not find a difference between monolinguals and L2 teachers, however they found a marginal advantage in short term memory for the interpreters. They reasoned that interpreters might have developed a different way of managing their languages, using other language control mechanisms than bilinguals. Another possible explanation for the lacking difference between interpreters and bilingual teachers might be provided by the Controlled Dose hypothesis (Paap 2018). According to this, an advantage in L2 might only surface while a new skill or the L2 is still being acquired. L2 teachers and interpreters might already have completed the formation and controlled execution stage of acquiring the L2 skill and moved on to the automatic execution stage. For the bilinguals and the interpreters, language control has become automatic and does not require much cognitive control. This might further explain why Woumans et al. (2015) found an advantage in their study, since their interpreting subjects were interpreting students and therefore possibly still at a stage where they worked on their L2 proficiency (van der Linden et al. 2018: 3–15).

Babcock & Vallesi (2017) compared professional interpreters to multilinguals in tasks tapping into short-term memory, working memory, conflict resolution and task switching. Results showed that interpreters performed better on verbal and spatial short term memory tasks than multilinguals and recalled more words in the verbal working memory task. Results did not show a difference between groups in conflict resolution and inhibition. In terms of the task switching performance, interpreters showed faster reaction times than multilinguals and smaller mixing costs for interpreters, indicating an advantage in sustained control for the interpreters but not in transient control (no group difference in switching costs). The authors emphasize that interpreters did not show enhanced executive functions beyond those seen in bilinguals but seemed to enhance functions that are explicitly related to the task of simultaneous interpreting (Babcock & Vallesi 2017: 406–415).

It has become evident from the studies reported so far that inhibition was least affected by an interpreter advantage in executive functions. Aparicio et al. (2017) specifically examined inhibition in interpreters and compared the performance of interpreters and highly proficient bilinguals in a language decision task and a bilingual Stroop task. They argued that in the language decision task, both active inhibition processes and overcoming inhibition processes are involved. When accessing the target word, the participant must inhibit the representations of the other language. When a switch trial occurred, participants had to overcome this inhibition of the previously inhibited representation. In the Stroop task, participants only have to apply active inhibition (interference suppression). The authors refer back to Green's ICM and Paradis' activation threshold hypothesis and explain that overcoming inhibition is more difficult than active inhibition and hypothesize that both groups should perform similarly when only active inhibition is required. When overcoming inhibition is required, the authors expected SIs to perform better than the bilinguals since simultaneous interpreting specifically trains switching and overcoming inhibition. In the language switch task, results showed an advantage for SIs, likely because bilinguals needed more cognitive resources to overcome inhibition than SIs. However, the task did not provide definite evidence that SIs really outperformed bilinguals in overcoming inhibition. The Stroop task provided more evidence. Although SIs responded faster than bilinguals, the analysis revealed no difference in terms of amplitude of the Stroop effect between groups. This indicates that the active inhibition process might not be influenced by simultaneous interpreting expertise and might be more related to a general proficiency of language use and frequency of language use. Furthermore, it suggests that the advantage of SI in the language decision task is due to an advantage in overcoming inhibition. It must be noted that the Stroop task although it does tap into executive functions was still linguistic in this case, making it more difficult to disentangle non-linguistic executive functioning and linguistic control. Still, results suggest that SIs possess an increased ability to deal with the interplay of activation and inhibition while keeping a task goal in mind (Aparicio et al. 2017: 1433–1446).

Zhong & Dong (2024) focussed more on the coordination skill of interpreting by implementing a dual-task situation and using ERPs to investigate lateralized readiness potentials (LRPs). Stimulus-locked LRPs reflect response-selection while response-locked stimuli reflect motor response and execution. The dual-task usually causes a reaction time cost, namely the difference in performance between the dual- and single-task condition. Smaller costs are connected to better coordination skills. The authors argue that interpreting involves frequent switches between overlapping subtasks such as listening and speaking or listening and note-taking in consecutive interpreting that need to be coordinated. In contrast to task switching, dual-tasking focusses more on how participants deal with overlapping tasks than on switching between tasks. In this study, the authors compared interpreting students to students of English. Prior to the dual task, participants were tested on working memory span and working memory updating; no significant differences between groups were found. Behavioural results showed that the interpreting students showed smaller dual-task costs than the control group which indicates an interpreter advantage in coordination. Stimulus-locked LRP ERP analysis showed smaller dual-task costs for stimulus-locked onset latency for the interpreting group than for the control group, again indicating a coordination advantage for the interpreters. Response-locked LRP onset latency analysis did not show a difference between groups. It seems therefore that interpreters possess better task coordination abilities than individuals who do not train this skill via interpreting training (Zhong & Dong 2024: 898–910). It has to be mentioned however that testing interpreters and controls on dual-tasking is not largely represented in literature in comparison to the “classic” executive function tasks.

The aforementioned studies all investigated a possible interpreter advantage in executive functions, likely brought about by the constant inhibition and activation of languages and the accompanied training of language control as well as executive control. However, these studies as well as other research in this field have yielded mixed results, as also summarised by Nour et al. (2020). A crucial drawback of the aforementioned research is that it is mostly comprised of cross-sectional designs (see Nour et al. 2020) that roughly group populations together as experts, novices or bilinguals without professional training and only provide insight into one point in time. However, as was established in Section 6.2, interpreting competence develops over time and it is difficult to infer competence or expertise from counting the number of years someone has spent as a simultaneous interpreter. Furthermore, longitudinal studies as well as more fine grained data gathering methods have the potential to offer a different insight into the underlying processes, however, neurolinguistic methods are rarely applied (Nour et al. 2020). Accordingly, a subsequent question that very few studies have investigated is how the acquisition of SI competence during training affects executive functions. The following paragraphs therefore present studies that applied mostly longitudinal designs and occasionally neurolinguistic measures to investigate the effect of SI training on executive functions:

As one of the first studies investigating this aspect of the interpreter advantage, Dong & Liu (2016) contrasted interpreting and translation. They tested consecutive interpreting students, translation students and Chinese students of English in the executive functions of inhibition (number Stroop), shifting (colour-shape switch task) and WM updating (n-back task). Applying a longitudinal design, they tested their participants at the beginning and at the end of one academic semester. In contrast to simultaneous interpreting, consecutive interpreting does not require the immediate overlap of comprehension and production. However, consecutive interpreting also requires immediate processing whereas more time is available during written translation. All three modalities have the frequent language switching in common. As indicated by the results, the interpreting experience yielded significant advantages in switch costs and updating skills. The translation experience only yielded marginally significant improvement in updating. None of the bilingual experiences resulted in any enhancement of inhibitory control or monitoring. The main difference between consecutive interpreting and written translation is the increased time pressure and the needs for immediate processing in interpreting. Consequently, the authors suggested that a high processing demand, as is brought about by interpreting training, might be a prerequisite for an enhancement of domain-general cognitive control (at least for switching and WM updating). To put it another way, a bilingual advantage tends to surface if the bilingual task at hand is cognitively demanding enough. Furthermore, Dong & Liu (2016) suggest the existence of a development curve for the enhancement of executive functions, proposing a steady development at the beginning of training followed by a plateau when cognitive functioning has reached a peak (Dong & Liu 2016: 3–10). This corresponds to current models of skill learning in the area of language as explained by Pliatsikas (2020) (see Section 3.1).

Babcock et al. (2017) as well as Rosiers et al. (2019) probed the question whether enhanced executive functions are acquired through selective training, or whether the advantage can be traced back to inherent characteristics that novice interpreters already possess at the beginning of their training. In their study, Babcock et al. (2017) compared interpreting students to translation students and students from non-linguistic fields within a longitudinal design. Participants were tested at the beginning and at the end of their MA studies in STM tests (letter span and matrix span), WM tests (o-span, symmetry span), the attention network task as well as a task switching paradigm. Results showed no advantages in memory or other executive function tasks for the interpreting students at the beginning of testing, indicating that interpreting students did not have an inherent advantage in EF. However, after the second round of testing, interpreting students showed substantial gains in verbal short-term memory compared to translation students and non-linguistic students, which was attributed to their interpreting training. No advantage in the other executive function tasks could be noted after the second round of testing. Concerning the lack of significant results for spatial short-term memory and working memory, the authors suggested that advantages in these categories might emerge with increasing interpreting experience after the training period of the MA (Babcock et al. 2017: 257–263). Rosiers et al. (2019) also tested possible pre-existing enhanced executive functions in interpreters, however not within a longitudinal design. They tested working memory capacity, inhibition, shifting and WM updating skills prior to the respective training of three experimental groups: MA interpreter students, MA translation students and MA multilingual communication students. The participants were administered the Simon task, the ANT, 2-back task and the colour-shape switch task. Furthermore, they had to complete a forward and a backward digit span task. The results did not indicate any significant difference between any of the groups in inhibition, shifting and updating. There was a small effect of working memory capacity indicated by a smaller span effect in the backward digit span task for interpreting students and multilingual students which was attributed to a possible inherent competence of students who opt for a course of studies with a specific oral component. Since no significant differences between the groups could be detected in the main executive function tasks, it can be concluded that interpreters did not possess an inherent executive function advantage before receiving interpreting training (Rosiers et al. 2019: 120–125).

Dong & Zhong (2017) investigated how interpreting experience affects executive functions in interpreting students with the scope of an ERP study. They tested two groups with more and less interpreting experience in the course of their university studies in a Flanker task. Results revealed that participants with more interpreting experience showed larger N1 and N2 amplitudes in congruent as well as incongruent trials, indicating a general advantage in attentional processing and conflict monitoring. Results for the P3 component were less conclusive since the first half of the P3 time window indicated a monitoring advantage and the second half of the P3 time window indicated an advantage in inhibition. In comparison to other studies on interpreting and executive functions this study suggests that interpreting as a form of extreme language control is a unique experience that differs from general bilingual processing (Dong & Zhong 2017: 196–202).

In one of the few longitudinal studies that apply neurolinguistic measures and also combine these with behavioural data, van de Putte et al. (2018) used fMRI to compare interpreting students and translation students at the beginning and at the end of a nine month training period in different executive functions tasks. They used the Simon task to investigate interference suppressions, a colour-shape switch task to tap into task switching as well as a verbal fluency task that included a language switching component to assess semantic fluency and language switching performance. Behavioural analysis did not reveal any significant effects of group or pre- or post-test on reaction time or accuracy. The authors explained that this lack of group difference could be attributed to the close relationship of interpreting and translating. Both tasks require formulating a source text into a target text and have some neural processes in common. Furthermore, it is possible that behavioural differences between groups might appear after a longer period of training or after expertise has been formed. Lastly, there is the possibility that both groups have already reached a ceiling in terms of behavioural advantages that can be derived from bilingual training and no additional boost could be reflected in reaction times or accuracy. Neural results provided further insights: A whole brain fMRI analysis revealed that compared to translation students, the interpreting students displayed higher involvement of the left superior temporal gyrus in the Simon task and an increased involvement of the right angular gyrus in the colour-shape switch task after the nine month training period. This can be interpreted as increased capacity of cognitive control, more specifically for interference suppression as well as inhibition aspects of switching between tasks and languages. In terms of structural connectivity, a subnetwork that connects the frontal regions with the basal ganglia that includes the left superior frontal gyrus, left/right medial superior frontal gyrus, left orbital superior frontal gyrus and the right pallidum showed increased structural connectivity for the interpreters compared to the translators. The right pallidum as a substructure of the basal ganglia seemed to be central to this network. The authors summarize that the basal ganglia in cooperation with frontal regions have been assumed to be implicated in executive functions, more specifically in suppressing and promoting responses that are in competition against each other and are governed by the frontal regions both in domain-general cognitive control as well as verbal control. Interpreting students showed increased connectivity compared to the translators in a second subnetwork that comprised the left SMA, right postcentral gyrus, right superior frontal gyrus, right middle temporal pole, right amygdala, the left inferior parietal gyrus and the superior parietal gyrus. Several of these areas, namely the cerebellum, the SMA and the parietal lobes overlap with the control network proposed by Green & Abutalebi (2013). In this network, the SMA is connected to conflict monitoring as well as initiation of speech in language switching while the parietal lobes are involved in maintaining task representations. The authors conclude that taken together, the results imply that the brains of trainee interpreters are subjected to neural changes in networks that are related to both executive control as well as language control (van de Putte et al. 2018: 247–255).

Zhao et al. (2024) hypothesized that consecutive interpreting students would solve conflict in a Flanker task better than non-interpreting students since they are more often confronted with interference from different languages. To investigate whether training the Flanker task would yield better results for the interpreter group than for the control group, the Flanker task was divided into two phases. For reaction times, results showed significant main effect with incongruent trials showing longer reaction times than incongruent trials and overall faster responses in the second half of trials but no main effect of group. However, interpreters showed lower error rates than non-interpreters. The ERP results showed significantly less negative N200 amplitudes for the interpreting group in both conditions than for the control group. Also, N200 amplitudes were less negative for interpreting students in the second half of the trials compared to the control group. The authors interpreted the results in favour of the interpreters needing less attention for monitoring while also making fewer errors compared to non-interpreters, presumably because of their enhanced monitoring skills due to training language control during interpreting (Zhao et al. 2024: 1–8).

As my review of the body of research concerning executive functions and SI shows, results are, although partially promising, still mixed, and inconsistent and do not provide definitive proof that SI enhances all executive functions. Among the three main executive functions, the evidence that interpreters show an advantage in working memory and working memory updating seems to be fairly consistent (Christoffels et al. (2006); Henrard & van Daele (2017); Rosiers et al. (2019); Babcock & Vallesi (2017); Macnamara & Conway (2014); Morales et al. (2015); Dong & Liu (2016)) with some exceptions (van der Linden et al. (2018); Babcock et al. (2017)). Task switching and flexibility also showed frequent positive results (Yudes et al. (2011); Macnamara & Conway (2014); Dong & Xie (2014); Dong & Liu (2016); Henrard & van Daele (2017); Babcock & Vallesi (2017)). Inhibition however isolated in a Simon, ANT or Flanker task rarely displayed an advantage and was mostly marked by null results (Yudes et al. (2011); van der Linden et al. (2018); Dong & Xie (2014); Dong & Liu (2016); Babcock et al. (2017); Babcock & Vallesi (2017); Rosiers et al. (2019); Aparicio et al. (2017)). Interestingly, the study by Dong & Zhong (2017) was able to show an advantage in inhibition using event-related potentials, while other studies testing interpreters on the same Flanker task but using behavioural measures could not detect an advantage in inhibition. Recent literature meta analyses such as by Nour et al. (2020), Hu & Fan (2021) and García et al. (2020) confirm this picture. Nour et al. (2020) only found an interpreter advantage in the executive functions of shifting and updating but not in inhibition. Interestingly, interpreters performed better than control groups in updating in cross-sectional studies, but studies did not show an improvement in updating for trainees in longitudinal designs. Contrary to this, shifting improved in both cross-sectional as well as longitudinal designs (Nour et al. 2020). Hu & Fan (2021) found significant evidence for an interpreter advantage in shifting, mixed findings for updating and only sparse evidence for an advantage in inhibition (Hu & Fan 2021: 157). In consideration of these mixed findings for both trainee interpreters as well as professional interpreters, García et al. (2020) argued that "[…] it would thus seem only specific functions, directly taxed in the exercise of SI, benefit from sustained practice of this activity" (García et al. 2020: 733). This means that some executive functions might just not be relevant to simultaneous interpreting or that the whole extent of an executive function is not affected by the enhanced processing demands of simultaneous interpreting. However, they highlight that pre- and post-training design studies such as by Babcock et al. (2017) are extremely valuable, since they have the capacity to show that if advantages occur, they are likely brought about by the simultaneous interpreting training. The interpreter advantage research field shares various shortcomings with the bilingual advantage research field. Sample sizes are low, although this is not uncommon in studies using methodologies from the cognitive sciences. Furthermore, participant information is not thoroughly reported, or interpreting experience and language proficiencies are not matched between groups. Tasks that assess executive functions vary greatly between studies, although some of the domains use validated tasks. The issue of unity and diversity of executive functions and their tasks remains. Finally, another topic is the granularity of measurement (García et al. 2020: 733–737). The majority of studies that explicitly examined executive functions in interpreters used reaction time and accuracy measures. Only a handful (Dong & Zhong (2017); Zhong & Dong (2024); van de Putte et al. (2018), Zhao et al. (2024)) use neurolinguistic measures, such as fMRI or EEG. However, as it has become clear in studies that examined general structural and functional differences between interpreters and other populations, these tools can give valuable insight that do not depend on the motor response of a button press to measure, for example reaction times.

As shown in Dong & Zhong (2017), more fine grained measurements such as electrophysiological techniques can provide a better insight into how the reaction of the brain changes if interpreting training is provided. What has become clear from this section and was also mentioned by García et al. (2020) is that there is still a lack in longitudinal studies while cross-sectional studies still dominate the research landscape. However, interpreting is an acquired skill that both includes the skilled management of two languages (and in extension of the executive functions associated with language control) as well as the successful implementation of interpreting specific processes, e.g. overlapping processes of comprehension and production. It is therefore extremely important to look at the development of a possible advantage in executive functions longitudinally while SI training progresses. As the end of Section 5.4 has shown, a renormalization of a skill might occur and there is a possibility that this renormalization or automation does not show itself in reaction times in a post-test, even if a longitudinal design was applied. Finally, as laid out in Section 6.1 by García (2014), it is not clear how soon after the beginning of training advantages in executive functions appear (García 2014: 232).

To sum up, this dissertation has so far shown how the brain changes in reaction to extraneous experience, how bilingualism as one of these extraneous experiences can reshape the brain both functionally and structurally in both linguistic as well as non-linguistic aspects, namely the domain-general executive functions. Different models of bilingual language identification and production have been discussed and with this, how executive functions play a role in language processing. Evidence for an overlap of executive functions and language control has been given from a behavioural as well as a neuroscientific point of view. The preceding sections have discussed simultaneous interpreting as an extreme form of language control, touching upon its characteristics, strategies applied in simultaneous interpreting, how interpreting competence emerges and how it is different from expertise and how executive functions are involved in both strategies as well as competence models. Furthermore, how simultaneous interpreting affects brain regions implicated in language control and executive control has been highlighted. Finally, as the last missing puzzle piece, studies examining the effect of simultaneous interpreting on executive functions have been examined specifically from both a behavioural as well as a neuroscientific point of view.

However, as discussed previously in this current section, studies that explicitly investigate how the emergence of SI competence through training affects executive functions and thereby laying the groundwork for the hypothesis that simultaneous interpreters possess enhanced executive functions at all, are extremely scarce, rarely longitudinal and mostly do not apply neuroscientific measures that would provide explicit insights into the workings of interpreters' brains. If longitudinal studies are applied, they are conducted in a pre- and post- test manner which does make it possible to determine whether there are any improvements in executive functions at the end of training. However, pre- and post-tests still mask what occurs between the two points of testing and do not provide an insight into the progression and possible normalization of executive functions as a result of training and competence acquisition. To this end, an empirical study that goes beyond the pre- and post-test methodology is conducted by using a triangulation of event-related potentials, behavioural data as well as competence data. The following Chapter 7 is dedicated to precisely outlining the research gap and formulating the hypothesis based on literature. Chapter 8 is dedicated to the methodology of the study and well as the results.

7. Research gap and hypothesis

Chapter 6 and Section 6.1 have shown that although different executive functions are mentioned in models of the interpreting process and models of interpreting competence (Gerver (2002); Setton (1999); Chabasse (2009); Seeber (2011)), they are not explicitly represented in interpreting process models. Models that look at the interpreting process in a more comprehensive way and from a neuroscientific perspective do mention brain areas that are implicated in executive functions (Paradis (1994); Fabbro (1999); Garcia (2019)). However, they also do not explicitly represent the contribution or involvement of executive functions in interpreting. From the structural, functional and electrophysiological evidence in Section 6.3 it has become clear that executive functions are likely to be implicated in the process of simultaneous interpreting and are affected by training language control and other interpreting specific functions (see e.g. Elmer et al. (2014); Becker et al. (2016); Hervais-Adelman et al. (2015b); Hervais-Adelman et al. (2017); Hervais-Adelman & Babcock (2020)). Some of the studies presented in Section 6.4 have shown an advantage or a change in executive functions for interpreters compared to non-interpreting controls or in pre- and post-tests among student interpreters (see e.g. Morales et al. (2015); Henrard & van Daele (2017)). However, evidence speaking for a change or an enhancement in executive functions due to professional simultaneous interpreting practice or simultaneous interpreting training is not consistent across all executive function tasks and studies (see e.g. Yudes et al. (2011); Dong & Xie (2014); Dong & Liu (2016); Woumans et al. (2015); Babcock & Vallesi (2017)) and in some studies is not present at all (e.g. van der Linden et al. (2018)).

It has become apparent that the research landscape around executive functions in interpreting still lacks both longitudinal studies, neurolinguistic methodologies as well as a more fine-grained insight into the process of interpreting training and the associated change in executive functioning. Many studies were purely behavioural (e.g. Morales et al. (2015); Macnamara & Conway (2014)), many studies only take the years of interpreting experience as an indicator for established interpreting competence, which does not reflect how much interpreting is done or how high the quality of the interpreting job really is. If studies investigated how interpreting training changes executive functions with neuroscientific methods, it was done within the scope of fMRI studies (see van de Putte et al. (2018)). Although fMRI provides important insights into structural and functional changes in the brain, its temporal resolution is quite low since it is a haemodynamic method. EEG studies, which would offer a precise insight into the temporal aspects of changes in executive functions, are extremely scarce (Dong & Zhong (2017); Zhong & Dong (2024)) and only investigate selected executive functions. There are no EEG studies that investigate how brain reaction and general executive function changes during interpreting training in a longitudinal design. If longitudinal designs were applied (e.g. Babcock et al. (2017)), this was done within a pre- and post-test methodology. Although pre- and post- measurements are a vital part of experimental testing and an established method in cognitive science, they may not cover the full extent of the SI training experience. As was laid out in chapters 2 and 3, the brain reacts to an initial confrontation with a new skill, and structurally and functionally adapts to the new challenges, especially in the realm of language (see e.g. Mechelli et al. (2004); Grogan et al. (2012); Hosoda et al. (2013)). However, as general models of skill learning (Chein & Schneider (2012)) and interpreting competence acquisition models have shown (Kutz (2010)), cognitive control shifts from a more controlled stage of skill execution to a more automatized phase of skill execution when competence increases. Within the field of language learning, this automation has been shown to result in a renormalization of previously emerged structural brain changes (Pliatsikas (2020)). The question is whether executive functions experience a similar renormalization in simultaneous interpreting after the simultaneous interpreting process has been automatized to a greater extent than in the beginning of training. If this is the case, some null results in changes of executive functions within pre- and post-test designs could be explained. Some studies have found changes in executive functions in student interpreters as early as after several weeks of training. The question therefore is at what exact point a change in executive functions occurs in interpreting students as an effect of training and whether this change lasts until the end of training, renormalizes or reaches a ceiling. It would therefore be crucial to take a more detailed look at how executive functions change throughout interpreting training that goes beyond testing at the beginning and at the end of training but also at intermediate intervals.

Lastly, none of the studies that examine executive functions have drawn conclusions for didactics. However, it would be insightful to know when interpreting training takes on a maximum effect and how interpreting competence, and to an extent executive functions, are affected by an interruption of training by for example an Erasmus semester.

Based on these gaps in research, the empirical part of this dissertation attempts to answer the question of when exactly changes in different executive functions take place in simultaneous interpreting trainees, how these changes progress and whether they reach a ceiling and are automized.

During SI, brain networks both implicated in linguistic processing as well as executive functions are involved in managing both languages that are used (see Hervais-Adelman & Babcock (2020) for a review); these networks correspond largely with the bilingual language control network put forth by Abutalebi & Green (2007; 2008) and Green & Abutalebi (2013). As several studies from the bilingual advantage research have shown, this overlapping network or language processing and executive control has the potential to result in enhancements of executive functions (see e.g. Bialystok et al. (2004); Costa et al. (2008); Kousaie & Phillips (2012); Morrison et al. (2019)). Simultaneous interpreting has been identified as a modality of extreme language control (see e.g. Hervais-Adelman & Babcock (2020); Hervais-Adelman et al. (2015b)) and several studies investigating a beneficial effect of SI on executive functions have partly found promising results (see e.g. Morales et al. (2015); Babcock & Vallesi (2017); Yudes et al. (2011)). Consequently, studies investigating the effect of SI training on executive functions have assumed that the enhanced control demands during SI training affect executive functions, and have partly provided evidence that an increase in SI competence through SI training results in structural brain changes (see van de Putte et al. (2018)) as well as a more efficient use of executive functions (see Zhao et al. (2024)).

The following study therefore investigates how progressing interpreting training affects the executive functions of simultaneous interpreting students. To answer this research question and shed a light on the temporal dynamics of changes in the brain, electroencephalography was chosen as the research method. The advantages of the EEG and event-related potentials are discussed in detail as part of the methodology in Section 8.1.2, where EEG studies connecting the relevant ERP components to the investigated executive functions are also presented.

Instead of a pre- and post-test design, interpreting students are tested at baseline before beginning their interpreting training and then at the end of every semester for four semesters which constitutes their whole MA Conference Interpreting. As done in previous studies such as by Dong & Liu (2016), Babcock et al. (2017), Rosiers et al. (2019) or van de Putte et al. (2018), non-interpreting translation students were used as a control group and were compared to the interpreting students at baseline level. Participants are tested on three commonly used executive function tasks: the Flanker task (see Eriksen & Eriksen (1974) and e.g. van der Linden et al. (2018)) to test inhibition, the colour-shape switch task (see Braver et al. (2003) and e.g. van de Putte et al. (2018)) to test task-switching performance and the n-back task (see Smith & Jonides (1997) and e.g. van der Linden et al. (2018)) to investigate working memory updating. Event-related potentials of the P300 (Sutton et al. (1965); Polich (2007); Daffner et al. (2011)) and N200 (Folstein & van Petten (2008); Cespón & Carreiras (2020)) ERP-components are examined. Interpreting competence is assessed by a validated questionnaire (TICQ) (Schaeffer et al. 2020) that participants fill out at the beginning and at the end of their interpreting training. A more detailed description of the study is given in the methodology in Section 8.1.

At baseline, neither the translator group nor the interpreter group has received prolonged interpreting training. Therefore, for the between group comparison at baseline level, I hypothesize that there will be no significant group difference in all tasks in terms of reaction times and accuracy, and, furthermore, that there will be no significant differences between groups in terms of voltage in the P300 and N200 time-windows of the respective tasks.

After baseline testing, only the interpreting group receives intensive interpreting training (see Section 8.1.3 for more information on workload). Therefore, for the longitudinal study of the interpreting group, I hypothesize the following:

Based on studies that showed gains in interpreting performance after training (see Tzou et al. (2012), Hervais-Adelman et al. (2015a), Yu & Dong (2022)), I hypothesize that the interpreting competence will increase from the start of the interpreting training to the end of the interpreting training as an effect of competence acquisition which will be mainly analysed by means of the TICQ-values (Schaeffer et al. 2020).

In terms of the between group comparison at baseline in the Flanker task, I hypothesize that there will be no difference between groups in terms of accuracy, reaction times and voltage in the respective ERP time-windows in any of the executive functions tasks based on the results by Babcock et al. (2017) and Dong & Liu (2016), who did not find a difference between interpreting group and two control groups in executive function tasks at the beginning of their longitudinal study. All participants should have the same level of knowledge and executive function ability prior to receiving their respective MA training in translation or simultaneous interpreting, so no inherent advantage of the interpreting group in executive functions should be present.

As previously shown in studies such as by Costa et al. (2009), Verreyt et al. (2016), Dong & Xie (2014) or van der Linden et al. (2018), I hypothesize incongruent trials producing longer reaction times and lower accuracy than congruent trials in the Flanker task, since they require more inhibition effort. Since Dong & Zhong (2017) found an effect of SI experience level on the Flanker interference effect, I hypothesize faster reaction times and increased accuracy over the course of SI training. For the EEG measures, I hypothesize based on literature such as by Kousaie & Phillips (2012), Wild-Wall et al. (2008), Johnstone et al. (2009) or Grundy et al. (2017b) that N200 amplitudes are larger (more negative) for incongruent trials compared to congruent trials in general since larger N200 has been associated with active suppression of irrelevant information (Kopp et al. 1996). I hypothesize that the N200 will become less negative for both conditions across the course of testing since increased SI experience through training has been associated with improved interference suppression (see Dong & Zhong (2017)) and the N200 has shown to be less negative when fewer resources for the suppression of irrelevant information need to be allocated (see Kopp et al. (1996), Cespón & Carreiras (2020)). Higher alpha power has been associated with increased inhibition (Klimesch et al. 2007). Therefore, for the frequency analysis, I hypothesize that alpha power in the alpha frequency range decreases over the course of testing, since the interpreting students will need to apply less inhibitory control to resolve the task.

For the colour-shape switch task I hypothesize longer reaction times and lower accuracy for switch compared to repeat trials (see Prior & Gollan (2011), Hernández et al. (2013), Garbin et al. (2010), Baene et al. (2015)) since they require more attentional control, interference suppression and updating of working memory in form of the stimulus-response mapping (see Jamadar et al. (2015), Cespón & Carreiras (2020), Polich (2007)). Based on results showing an improvement in RT and accuracy after SI training (see Dong & Liu (2016) or Babcock et al. (2017)), I hypothesize faster reaction times and increased accuracy over the course of testing.

Furthermore, I hypothesize larger N200 (more negative) amplitudes for switch trials relative to non-switch trials (see e.g. Timmer et al. (2017), López Zunini et al. (2019), Declerck et al. (2021)) since they require more interference suppression from the task rule of the previous trials and more resource allocation to interference suppression has been associated with more negative N200 amplitudes (see Kopp et al. (1996), Cespón & Carreiras (2020)). Again, I hypothesize that the N200 will become less negative over the course of testing for both conditions since interpreting training been associated with improved interference suppression (see Dong & Zhong (2017)) and the N200 has shown to be less negative when fewer resources for inhibitory effort need to be allocated amplitudes (see Kopp et al. (1996), Cespón & Carreiras (2020)).

For the P300 amplitude, I hypothesize smaller P300 amplitudes for switch than for repeat trials based on task-switching studies such as by López Zunini et al. (2019), Periáñez & Barceló (2009) or Chen et al. (2022) (see Cespón & Carreiras (2020) for a review), since smaller P300 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). I hypothesize that over the course of testing, P300 amplitudes increase for both conditions since SI training has shown to improve task switching abilities (see Dong & Liu (2016), Babcock et al. (2017)) rooted in attention switching (Monsell 2003), and allocating fewer cognitive resources to attention switching and stimulus categorization has been associated with increased P300 amplitudes (Cespón & Carreiras 2020).

In the frequency domain, I hypothesize less alpha power over the course of testing in the N200 time-window as well as in the P300 time-window at the respective fronto-central and parietal electrode, since alpha power has been associated with increased inhibition processes (Klimesch et al. 2007) and I expect the inhibitory effort to decrease with progressing SI training (see Dong & Zhong (2017)).

For the 2-back and 3-back tasks, I hypothesize lower accuracy for match trials than for no-match trials as demonstrated e.g. in van der Linden et al. (2018) and Szmalec et al. (2011). Accordingly, as previously found in Szmalec et al. (2011) and Barker & Bialystok (2019), I hypothesize faster reaction times for match trials than for no-match trials. Consequently, I hypothesize faster reaction times over the course of interpreting training as demonstrated in Dong & Liu (2016) and higher accuracy over the course of interpreting training as demonstrated in a numerical increase of accuracy in Babcock et al. (2017). However, I hypothesize significantly slower reaction times and lower accuracy for 3-back match trials than for 2-back match trials since reaction times and accuracy have been shown to decrease with increasing task difficulty (see Morrison et al. (2019), Morales et al. (2015), Daffner et al. (2011)), and the processing load on updating the working memory increases with increasing load conditions (Jaeggi et al. 2010). In terms of electrophysiological results, I hypothesize larger P300 amplitudes for match trials than for no-match trials for the 2-back and 3-back task since categorizing the stimulus is more effortful in the no-match than in the match condition (see Barker & Bialystok (2019)). I furthermore hypothesize smaller P300 amplitudes in the 3-back than in the 2-back task since P300 has been shown to decrease with increasing memory load as in Dong et al. (2015) or Morrison & Taler (2020). Lastly, I hypothesize that the P300 amplitude will increase across the time of testing since studies such as by Dong & Liu (2016) have shown behavioural improvements in WM-updating after SI training and available cognitive resources for working memory updating have shown to result in larger P300 amplitudes (Morrison & Taler 2020).

For the frequency analysis, I hypothesize less alpha power in the P300 time-window over the course of testing, since participants will benefit from SI training and become more routinised in updating their working memory and categorizing stimuli and therefore require less inhibitory control (Klimesch et al. 2007).

Finally, based on models of general and linguistic skill acquisition (see Chein & Schneider 2012, Lövdén et al. 2013, Pliatsikas 2020), the controlled dose hypothesis by Paap (2018) as well as speculations from studies such as by Dong & Liu (2016) who specifically investigated the effect of interpreting training on executive functions, I hypothesize a ceiling effect for both behavioural and electrophysiological measures that reflects an automation of the respective executive functions involved during SI training to accommodate the increased cognitive control demands during SI training.

8. Study

The goal of the study is to investigate how executive functions are affected by interpreting training over time, at which point interpreting training has the most impact on executive functions and how the effect of training on executive functions progresses over the whole course of the MA Conference Interpreting. In contrast to most longitudinal studies, no pre- and post-test methodology was used but instead a more faceted approach with testing after every semester. For this purpose, a longitudinal design was applied, and electroencephalography (EEG) was selected as the data gathering method of choice to gain direct insight into the temporal dynamics of longitudinal changes of the reaction of the brain to executive functions-related stimuli. The goal is to determine if executive functions and interpreting competence acquisition progress during interpreting training and whether a ceiling effect or a normalization can be observed at some point.

8.1. Methodology

The following chapters describe the methodology of the study, i.e. the tasks that are used, which data gathering method was chosen, the selected stimuli, which participant groups took part in the study as well as the apparatus and procedure. Finally, the data analysis process is explained, and the results of the study are presented.

8.1.1. Tasks

The present longitudinal study was conducted to explore the impact of interpreting training on functional plasticity by focussing on the changes in domain-general executive functions. More specifically, tasks were chosen that target attentional processing, interference suppression, task switching (mental flexibility) and working memory updating.

To target interference suppression, a modified version of the Eriksen Flanker task was applied (Eriksen & Eriksen 1974). The Eriksen Flanker task has been widely used in cognitive psychology to test inhibitory functions, more specifically interference suppression. In the cognitive sciences, there are other tasks that also target the inhibitory executive function such as the go/no-go task or the Stroop task. However, as explained by Bunge et al. (2002a) and mentioned in Chapter 4.1, there are two kinds of inhibitory control; interference suppression that is investigated by means of the Flanker task and response inhibition that is investigated by go/no-go paradigms or Stroop tasks. Since Luk et al. (2010) have shown that bilingualism affects neural mechanisms for interference suppression rather than for response inhibition, and the Flanker task and its variations have been widely used to investigate the advantage in inhibitory control in bilinguals (see e.g. Costa et al. (2008), Kousaie & Phillips (2012), Verreyt et al. (2016)), the Flanker task was the task of choice for this experimental manipulation.

Although versions of different complexity exist, the Flanker task typically consists of congruent and incongruent stimuli, mostly arrows or letters. As described in Chapter 5.4, the Flanker task has also been frequently applied to test interference suppression or inhibition in bilinguals and multilinguals and compare the results to monolinguals (e.g. Kousaie & Phillips (2012); Kousaie & Phillips (2017); Markiewicz et al. (2023); Luk et al. (2010)). Furthermore, as laid out in Chapter 6.3, several studies have applied the Flanker task to compare interference suppression in simultaneous interpreters and bilinguals, translators or monolinguals (e.g. Dong & Xie (2014); van der Linden et al. (2018); Zhao et al. (2024); Dong & Zhong (2017)). The incongruent condition is cognitively more demanding, since participants must suppress or inhibit the interfering components of the incongruent stimulus to make a correct decision. Therefore, reaction times in the incongruent condition are usually longer than in the congruent condition (Luk et al. (2010); Dong & Xie (2014)). In terms of ERP, the Flanker task as a task representing interference suppression or inhibition, is usually associated with the N200 component (e.g. Heil et al. (2000); Wild-Wall et al. (2008)). As laid out in Chapter 8.1.1, the incongruent condition in the Flanker task is usually associated with larger N200 amplitudes and later N200 latencies compared to the congruent condition (e.g. Purmann et al. (2011); Kousaie & Phillips (2012); Kousaie & Phillips (2017)). In the present study, the Flanker task consists of a row of 5 arrowheads either all pointing in the same direction (congruent condition) or a row of 5 arrowheads in which the central arrow is pointing in a different direction than the other 4 (incongruent condition). Participants are asked to indicate the direction of the central arrow on a keyboard, using the “J” key to indicate the middle arrowhead pointing to the right and using the “F” key to indicate that the middle arrowhead points to the left. Chapter 9.1.5. contains a detailed description of the stimuli adaption.

To investigate the impact of interpreting training on task switching or mental flexibility, a colour-shape switch task was used. Among task-switching paradigms, the colour-shape switch task has widely been applied in examining the bilingual advantage in task switching (see e.g. Dong & Liu (2016), López Zunini et al. (2019), Garbin et al. (2010)). Furthermore, studies investigating the influence of SI on task switching exclusively use the colour-shape switch task (see Becker et al. (2016), Dong & Liu (2016), Babcock et al. (2017), Babcock & Vallesi (2017), Rosiers et al. (2019)).

As described in Chapter 4.2, colour-shape switch paradigms usually consist of two conditions “colour” and “shape” that have two levels each. In the case of the present study, the colours red and blue as well as the shapes circle and triangle are used. All conditions and levels are mapped on a keyboard. Additionally, colour-shape switch tasks consist of two different types of cues that indicate the task that is to be performed on the following stimulus. In the present study, a “shape-cue” consisting of three 2-D shapes and a “colour-cue” consisting of a spiral shaped colour gradient were used. As explained in Chapter 4.2, task switching paradigms are made up of “repeat-trials” where the same task rule needs to be followed as in the previous trial and “switch-trials” which warrant for a switch from the previous task rule to a different task rule indicated by the cue. During the experiment, participants initially see a fixation cross, followed by one of two cues indicating the attribute they need to react to with the keyboard. The cue is then followed by either a circle or a triangle in blue or red. The participants had to use their left hand and the keys “S” and “D” to react to the shapes and their right hand and the keys “K” and “L” to react to the colours. Switching costs and mixing costs are then calculated, with mixing costs indicating the performance costs on repeat trials in mixed-task blocks and trials in the single-task blocks and switching costs indicating the performance costs on switch trials vs. repeat trials only in the mixed-task block (see also Chapter 4.2). Colour-shape switch tasks have frequently been applied to test the executive functions of shifting or mental flexibility in bilinguals (e.g. Prior & Gollan (2011); Garbin et al. (2010)) as well as simultaneous interpreters (e.g. Dong & Liu (2016); van de Putte et al. (2018); Rosiers et al. (2019); Babcock et al. (2017); Babcock & Vallesi (2017)). The switch condition generally requires more effort than the repeat condition, which is why reaction times are usually longer in switch conditions than in repeat conditions (e.g. Prior & Gollan (2011); Garbin et al. (2010); Babcock & Vallesi (2017)). In ERP studies applying the colour-shape switch task, the P300 component and the N200 component were associated with allocation of attentional resources during task switching (P300) as well as interference suppression from previous trials (N200). The N200 amplitude is usually larger in switch trials compared to repeat trials and the P300 amplitude is usually smaller in switch trials than in repeat trials (see Jamadar et al. (2015) for a review). A detailed description of the stimuli used in the present study follows in Chapter 9.1.5.

To target working memory updating, the n-back task (Smith & Jonides 1997) was applied since it is widely used when investigating bilingual advantages in working memory updating (Hansen et al. (2016), Morrison et al. (2019), Morales et al. (2015), Barker & Bialystok (2019), Comishen & Bialystok (2021)) as well as interpreter advantages in working memory updating (Dong & Liu (2016), van der Linden et al. (2018), Rosiers et al. (2019)). Furthermore, as explained by Hansen et al. (2016), in contrast to other tasks that tap into working memory such as the o-span task, the executive load of the n-back task is higher, specifically regarding continually refreshing items held in working memory and managing interference from currently irrelevant but previously relevant items (Hansen et al. 2016: 56). This fits well with the high cognitive load during simultaneous interpreting, making the n-back task the task of choice in this dissertation.

The n-back task usually consists of different difficulty levels and contains a continuous stream of stimuli (Jaeggi et al. 2010: 394). In the present study, the 2-back and 3-back

versions were applied. In the 2-back task, participants see a sequence of letters on a screen and must press a button if a letter matches the letter that was shown two trials before. Respectively, in the 3-back version, participants must match a target stimulus with the stimulus that was shown three trials before. As explained by Jaeggi et al. (2010), resolving the conflict created by matching a stimulus to the previous one requires inhibitory control. Furthermore, participants have to consistently update their working memory as described in Polich (2007). Participants need to keep the representation of the relevant stimulus in working memory. When a match is detected and the next stimulus becomes relevant, participants need to update their working memory. Some n-back tasks also contain lure-trials to increase task difficulty, such as Szmalec et al. (2011). The n-back task has been used to determine working memory updating proficiency in bilinguals (e.g. Hansen et al. (2016)) as explained in Chapter 5.4 as well as in simultaneous interpreters (e.g. van der Linden et al. (2018); Rosiers et al. (2019); Dong & Liu (2016)) as explained in Chapter 6.3. Reaction times for match trials have been reported to be faster than reaction times for no-match trials (e.g. Szmalec et al. (2011)). When using ERP in conjunction with an n-back task, the P300 component was usually examined since the P300 has been related to stimulus categorization, working memory updating and working memory load(see Chapter 9.1.1). The P300 amplitude is usually smaller as working memory updating load is high or increases (see Cespón & Carreiras (2020) for a review)
A detailed description of the stimuli used in the n-back task follows in Chapter 9.1.4.

To motivate my decision to use EEG to investigate the process of executive functions during interpreting training, the following chapter describes the advantages and caveats of the EEG in detail and elucidates on the event-related potential components that are relevant to this study.

8.1.2 EEG and ERPs

Research applying electroencephalography began with Hans Berger's ground-breaking experiments in 1929, in which he observed that, by placing an electrode on the scalp, it was possible to record electrical activity of the brain caused by active neurons as voltage on the scalp. This signal was then amplified and changes in voltage over time could be mapped. This technique shaped the scientific and clinical research of the following decades (Luck 2005: 3–4).

However, in its unprocessed form, EEG data is a highly elusive way to measure brain activity since it contains a mixture of hundreds of different sources of neural activity. One technique to draw conclusions from an EEG signal is the so called event related potential (ERP) technique.

ERPs are defined as *"[…] voltage fluctuations in the ongoing electroencephalogram (EEG) that are time-locked to an event, such as the onset of a stimulus or the execution of a manual response"* (Luck & Kappenmann 2012: 5)". These neural reactions can be extracted from the entirety of the EEG signal by for example different averaging techniques (Luck 2005: 4). The neural basis of the EEG signal lies in the post-synaptic potential of pyramidal cells. As an action potential reaches the presynaptic membrane, neurotransmitters are released in the synaptic cleft and bind with receptors in the postsynaptic cell membrane. In response, ion channels open or close and ions flow in and out of the cell which causes a change in electrical potential: the post-synaptic potential (Luck 2005: 28 et seqq.). ERPs therefore do not reflect action potentials. In order to record changes in scalp-recorded voltage, post-synaptic potentials need to occur simultaneously in a large number of neurons that need to be spatially aligned perpendicularly to the cortical surface, which is mostly the case in groups of so called pyramidal cells in the cerebral cortex. When a post-synaptic potential occurs, polarity changes occur in the extracellular space around the pyramid cell and inside the pyramid cell, creating a so called dipole. Dipoles have positive charges at one and negative charges on the other end. The dipoles can add together to form one large dipole called an equivalent current dipole and then be recorded on the scalp surface if there are enough spatially aligned dipoles when the neurons are simultaneously active (Luck & Kappenmann 2012: 5). If the neurons do not fire together, the signal is too weak to be detected by the electrode. Further difficulty is caused by the fold of the cortex. When dipoles are arranged in more than 90 degrees from each other, they are likely to cancel each other out (Luck 2005: 31).

In comparison to the overall EEG signal, ERPs are very small. They only become visible as a result of combining multiple EEG epochs together, which then form an averaged ERP waveform (Luck & Kappenmann 2012: 5). The waveforms show a sequence of positive and negative voltage deflections called peaks that "reflect the sum of several relatively independent underlying latent components" (Luck 2005: 51). It is crucial to differentiate between the peaks in the ERP waveform and the so called ERP components that are embedded in the ERP waveform (Luck 2005: 51). Luck (2005) defines these ERP components as *"Scalp-recorded neural activity that is generated in a given neuroanatomical module when a specific computational operation is performed"* (Luck 2005: 59) and Luck & Kappenmann (2012) add that they can be considered as *"scalp-recorded voltage change that reflects a specific neural or psychological process"* (Luck & Kappenmann 2012: 4). It is important to acknowledge that the peaks are not completely congruent with distinct ERP components. Isolating ERP components is an ongoing issue. When neural processes unfold in the brain, different groups of neurons or in other words up to ten equivalent current dipoles can be active simultaneously in different brain regions. If ERP components are related to specific neural processes, every equivalent current dipole can be seen as a separate ERP component that occurs simultaneously with other ERP components. The ERP waveform from a specific electrode mirrors the contribution of different ERP components that are simultaneously active, which can lead to a superposition of different ERP components (Luck & Kappenmann 2012: 6–8). In order to separate the observed ERP waveforms from the underlying ERP components, approaches such as principal component analysis (PCA) or independent component analysis (ICA) can be applied (Luck 2005: 58).

A scalp recorded EEG is a non-invasive technique, which gives the EEG advantages over other forms of collecting brain data. ERP have an excellent temporal resolution of 1ms or better which gives them an advantage over fMRI data, since fMRI is based on the haemodynamic response that extends over seconds. However, the ERP technique has low spatial resolution. The voltage that is recorded at a single electrode at a certain event time-point is a summation of activity from various different sources (Luck 2005: 25–26). An electrode placed on a left frontal region of the scalp could therefore also pick up on activity that occurs more towards the midline. This caveat

has been described as the “single greatest shortcoming of the ERP technique” (Luck 2005: 26).
ERP components are usually named according to their polarity, P for a positive deflection of the electrical signal and N for a negative deflection. Additionally, they are named according to the time it takes on average in milliseconds for the ERP component to appear after the stimulus was presented, e.g. P300 or N200 (Banich & Compton 2018: 50).

The ERP approach of analysing the EEG signal lies in the time-domain, where many trials are epoched and time-locked to an event which results in the ERP waveform. Time-frequency analysis lies in the frequency domain since it focusses on how the frequency power spectrum changes when EEG data is time-locked to an event (Makeig et al. 2004: 204). The EEG signal contains different frequency bands, from very low frequencies such as theta at 4 Hz over alpha frequencies at c.a. 10 Hz to beta frequencies at c.a. 15Hz and gamma frequencies (40 Hz) (Banich & Compton 2018: 85). Alpha has been defined as the “[…] dominant brain wave rhythm in the normal human EEG and consists of oscillations in the 8-13 waves per second range” (Antonenko et al. 2010: 429–430). Waves per second is referred to as Hertz (Hz). Changes in the alpha power band have been observed in reaction to different events. The most conventional alpha response has already been observed by Hans Berger in the beginning of EEG research. Alpha waves were observed when participants had their eyes closed, however, upon opening of the eyes, alpha power was suppressed. Contrary to these findings, alpha suppression has also been observed during task performance and is therefore not only associated with opening the eyes to process visual stimuli (Klimesch et al. 2007: 65). Klimesch et al. (2007) interpret alpha suppression in terms of desynchronization (Klimesch et al. 2007: 65). Event-related desynchronization (ERD) is the “decrease in power of the rhythmic activity within the alpha band” (Pfurtscheller 1977: 757) and describes a decrease in synchrony of neurons involved in a neural process. While ERPs are thought to reflect how cortical neurons respond to changes in stimuli, ERD are thought to reflect how the activity of local interactions between neurons changes, indicated by different frequency bands (Pfurtscheller & Da Lopes Silva 1999: 1842–1843). Contrary to ERD, event-related synchronization (ERS) represents the concept of "[…] increase in synchrony of the underlying neuronal population […]" (Pfurtscheller & Da Lopes Silva 1999: 1842). Klimesch et al. (2007) have related ERD and ERS in the alpha band with different cognitive states. They propose that ERD as a decrease in alpha band power reflects active cognitive processing and therefore excitatory brain processes. In contrast, they propose that ERS, an increase in alpha power, mirrors inhibition of some cognitive processes. They define inhibition as a top-down control process and describe top-down processes as central executive control functions that are related to attentional control to keep processes focussed. Inhibition is used to prevent interferences. This definition of inhibition and executive functions corresponds to the definition used for this dissertation in Chapter 4.1. They conclude that executive control of attention is linked to increased alpha activity (Klimesch et al. 2007: 65–69).

On the basis of ERD, Makeig (1993) introduced the analysis of event-related spectral perturbation (ERSP) to uncover event-related aspects which the ERP averaged signal does not contain. It is grounded in the concept of ERD but measures the changes in amplitude of the whole EEG frequency spectrum in response to an event, taking into account multiple frequency bands which offers more information than ERD (Makeig 1993: 284). This technique issues a two-dimensional ERSP image that shows the mean change in spectral power from baseline in decibel (Makeig et al. 2004: 204). According to Makeig et al. (2004), ERSP and ERP are "nearly complementary" (Makeig et al. 2004: 204), since the ERSP changes cannot be well represented in the time-oriented features of the ERP, and ERPs cannot be well represented within the frequency domain. Therefore, to get a comprehensive overview of my data, I chose to also look at the time frequency patterns of selected participants.

As also described by Cespón & Carreiras (2020), when event-related potentials are used to research executive functions in bilinguals, analysis almost exclusively focuses on the N200 component to examine interference suppression and the P300 component to examine attentional switching and working memory tasks (see e.g. Markiewicz et al. (2023), Morrison et al. (2019), Barker & Bialystok (2019), Kousaie & Phillips (2012, 2017), López Zunini et al. (2019)). The same is valid for research on the interpreter advantage, although there are extremely few studies that apply ERPs (see Dong & Zhong (2017), Zhao et al. (2024)). An important aspect about the N200 and the P300 components is that they are well characterized ERP components that have been studied for a long time, also in the context of the bilingual advantage in executive functions (see above). Since relevant ERP literature such as by Luck (2005) recommend focussing on only one or two ERP components per experiment and to use well-studied ERP components (Luck 2005: 62), the choice to examine the N200 and P300 component in the context of this study to investigate how the brain reaction of SI trainees in different executive functions tasks change through SI training is justified.

The two ERP components will be described in more detail in the context of executive function tasks in the following paragraphs.

The N200 component can be described as a second negative peak occurring in the averaged ERP waveform. It often has a fronto-central distribution and usually peaks between 200-350ms after the stimulus has been presented (Cespón & Carreiras 2020: 316). Although sometimes found in auditory stimuli, it is mostly elicited by visual stimuli. As proposed by Folstein & van Petten (2008), the N2 should be divided into several subcomponents: a fronto-central component associated with mismatch and novelty detection, a fronto-central component associated with cognitive control that especially attends to response inhibition, conflict resolution and monitoring and two posterior N2 components associated with visual attention. Studies investigated the N2 component more thoroughly for the first time in the late 1970s and early 80s. Back then it was divided into the subcomponents N2a and N2b, the latter often being observed with P3a component in oddball studies. Later however, a third subcomponent, the N2c was observed in combination with P3b. N2a is nowadays often called mismatch negativity (MMN) for auditory modalities, and visual MMN for visual correlates (Folstein & van Petten 2008: 152–153). The N2 was especially investigated in the context of go/no-go paradigms, where participants either had to react to “go” stimuli or refrain from reacting in “no-go” trials, where a prepotent response needs to be overridden.

There is little doubt that N2 represents cognitive control processes and executive functioning (see Folstein & van Petten (2008)). However, there has been a lot of debate about whether the N2 represents conflict monitoring and therefore primarily engages the ACC, or whether it rather represents inhibition processes (see Nieuwenhuis et al. (2003) and Pfefferbaum et al. (1985)). Some researchers suggest that the N2 only represents the monitoring of response conflict. Especially Nieuwenhuis et al. (2003) speak in favour of the conflict monitoring hypothesis. According to this, an increased N2 amplitude in no-go trials reflects an electrophysiological correlate of conflict monitoring in the ACC and not inhibitory processes. They used a go/no-go paradigm in which they varied the frequency of go and no-go trials (20%, 50%, 80%) as well as inverse dipole modelling to find the neural source of the N2. They observed an N2 in go trials even if they were rare, which they interpreted against the inhibition hypothesis since no inhibition should be required in go trials, only monitoring. The results of the inverse dipole modelling pointed towards the ACC as N2 source (Nieuwenhuis et al. 2003: 24).

However, there are many accounts of research that support the hypothesis that N2 represents inhibition processes or at least, not purely monitoring processes. One of the first research accounts for this was presented by Pfefferbaum et al. (1985). They used a go/no-go task with an equiprobable amount of go and no-go trials as well as semantic and non-semantic stimuli. Furthermore, participants had to give an overt motor response in from of a button press in one condition and only count go trials but not no-go trials in their head in another. As a result, the N2 did not seem to be limited to overt motor responses and also occurred in the silent counting condition. Furthermore, no-go stimuli yielded a larger N2 than go stimuli. Since both stimulus types were equiprobable, this effect cannot be attributed to frequency of no-go trials, only to their relevance. They concluded that the N2 reflects frontal inhibitory mechanisms that come into action on no-go trials. This frontal N2 effect that is larger in no-go trials was found by various researchers and was explained in favour of a frontal inhibition mechanism (Jodo & Kayama 1992; Falkenstein et al. 1999; Dong et al. 2009; Heil et al. 2000). Bokura et al. (2001) used a go/no-go paradigm with a low no-go probability to examine how N2 is represented in the brain using low-resolution electromagnetic tomography. Results indicated that the no-go N2 component stems from the right lateral orbifrontal cortex as well as the ACC (Bokura et al. 2001: 2225–

2231). Lavric et al. (2004) investigated a similar question, but matched go and no-go responses. They hypothesized that the no-go N2 would reflect inhibitory processes. They found higher activation in the ventromedial prefrontal cortex (vPFC) as well as the dorsolateral prefrontal cortex (DLPFC) in the N2 time window, not in the ACC. They inferred that the N2 component possibly has two neural substrates, the ACC for conflict monitoring and the vPFC/DLPFC for inhibition. They postulated that when stimuli are matched for probability, N2 will reflect inhibition processes (Lavric et al. 2004: 2484–2487).

Although it has become clear that the N200 amplitude is consistently larger on conflict trials as opposed to non-conflict trials, it is important to discuss the impact of processing demands on the N2 component in the context of this dissertation. Jodo & Kayama (1992) provided first insights into this topic when they altered the time-period that participants were given to respond to a stimulus between groups in a go/no-go task. As frequently observed, results showed that N200 amplitude was significantly more negative in no-go trials than in go trials, however, N2 amplitude was also more negative for the group that was given less time to respond to a stimulus than for the group that was given more time to respond to a stimulus. The results suggested an association between larger (or more negative) N2 amplitudes and longer latencies in no-go trials and increased effort to withhold a response, corresponding to an increased activation of the inhibitory control system of the brain (Jodo & Kayama 1992: 480–481). This would mean that the N200 is modulated by how much effort needs to be applied in suppressing interference in incongruent trials or suppressing the false response option. It could be argued therefore that the less inhibition effort is needed, the less negative the N200 should become.

The N2 has not only been investigated in the context of go/no-go paradigms, but also in the context of Flanker tasks which also tap into the function of inhibition or, more precisely, interference suppression as a form of inhibition (see Chapter 4.2). This is relevant for this dissertation, where I used the Flanker task (see Chapter 9.1.3) to observe changes in executive functioning during the development of interpreting competence. In the context of the Flanker task, which contains congruent and incongruent conditions, incongruent trials have shown to elicit larger N2 responses compared to congruent trials (Wild-Wall et al. 2008; Heil et al. 2000; Johnstone et al.

2009; Kousaie & Phillips 2012, 2017; Dong & Zhong 2017; Markiewicz et al. 2023). This can be traced back to the active suppression or inhibition of the irrelevant information of the incongruent flankers. Specifically in terms of the Flanker task, Kopp et al. (1996) summarized that the N200 might reflect the inhibitory activity that comes up when a false response option is automatically primed. That would mean that in incongruent trials, a false response option is coactivated but then inhibited by exerting executive control (Kopp et al. 1996: 284). This corresponds to the coactivation of both languages in the mental lexicon described in chapters 5 and 5.1 and the inhibition that is needed to suppress the non-target language during simultaneous interpreting. As summarized by Cespón & Carreiras (2020), increased N200 amplitudes are assumed to reflect increased cognitive effort and more resource allocation to interference suppression (Cespón & Carreiras 2020: 317).
In terms of N200 latency, earlier peak latencies have been observed for congruent compared to incongruent trials in a Stroop task and a Simon task and earlier peak latencies have been observed for bilingual older adults compared to monolingual older adults in a Stroop task and in a Flanker task (Kousaie & Phillips 2017: 29). In their review, Cespón & Carreiras (2020) hypothesized that bilinguals should display faster N200 latencies than monolinguals in conflict trials that require inhibition processes since they are more trained in inhibition due to their experiences with language control. However, they review that so far, faster N200 latencies have only been shown in older adults and children and not in young adults (Cespón & Carreiras 2020: 320–323).

The Flanker task or modified forms of the Flanker task are frequently used in the context of bilingual advantage research to investigate executive functions in bilinguals compared to monolinguals (see e.g. Costa et al. (2008), Emmorey et al. (2008), Paap & Greenberg (2013), Grundy et al. (2017b), Markiewicz et al. (2023)). As already laid out in Chapter 5.4, the bilingual advantage hypothesis relies on the premises that two language system are activated non-selectively and are active at the same time, and that language control is needed to prevent languages from interfering. These language control processes are a part of cognitive control and similar to general executive control processes. Constantly managing two languages in bilinguals is therefore thought to train the cognitive control mechanisms and changing general cognitive control abilities, which also include conflict monitoring and general inhibitory processes. In the case of the Flanker task, inhibition is taxed in the form of interference suppression.

Consequently, various studies have investigated the effects of bilingualism on conflict monitoring and inhibitory processes in bilinguals and monolinguals by analysing the N200 component. Kousaie & Phillips (2012) used among other tasks a Flanker task and compared monolinguals and bilinguals. Results yielded larger N200 amplitudes for incongruent trials than for congruent trials in both conditions, however, there was no significant group difference. They only found a group difference in terms of error-related processing. Kousaie & Phillips (2017) replicated this study with older bilinguals and found that the N200 amplitude peaked earlier in bilinguals than in monolinguals in incongruent trials in the Flanker task. They concluded that the bilinguals were faster at conflict monitoring than monolinguals on incongruent trials. Some studies have also investigated the effect of low conflict and high conflict ratio on the N200. Grundy et al. (2017b) used the Flanker task and the N200 component to investigate whether bilinguals are quicker in switching attention away from irrelevant information than monolinguals by testing the effect of previous trial congruency on the N200 component in a Flanker task. They calculated the sequential congruency effect (SCE) which indexes how performance adjusts depending on the congruency of the previous trial. They did not find a main effect of language group on the mean N200 amplitude, but the N200 amplitude was significantly larger for incongruent trials than for congruent trials. Furthermore, they found that when the previous trial was incongruent, the SCE of the Flanker was significant for monolinguals, however, for bilinguals, the SCE was smaller than for monolinguals and only reached significance in some cases. The authors suggested that bilinguals were able to disengage attention more rapidly from the previous trial in the critical incongruent trials since they have trained to do so due to their experience in managing multiple languages (Grundy et al. 2017b: 43–52). Purmann et al. (2011) examined whether there is an adaption of monitoring behaviour in Flanker tasks with frequent conflict trials relative to non-frequent conflict trials. They found that although the N2 amplitude was bigger for incongruent than for congruent trials in both frequent and infrequent stimuli conditions, the N200 amplitude was smaller when conflict was frequent than when conflict was rare (Purmann et al. 2011: 56). As already mentioned in Chapter 5.4, Markiewicz et al. (2023) compared monolinguals and bilinguals on the Flanker task with a incongruent/congruent ratio of 25%/75% and found that N200 was larger (more negative) for bilinguals than for monolinguals in both conditions and a P300 that was smaller (less positive) for bilinguals than for monolinguals. This was interpreted as evidence that incongruent trials created greater conflict monitoring demands and that bilinguals allocated more resources to monitoring than monolinguals and also resolved conflict earlier (Markiewicz et al. 2023: 133–143).

In the context of simultaneous interpreting as a form of extreme language control, only a handful of studies have used the ERP technique and specifically the N200 component to examine inhibitory control in interpreters. According to the rationale built up in this dissertation, simultaneous interpreters should possess enhanced inhibition skills since they have trained their language control and more specifically the ability to inhibit the respective non-target language when listening and speaking. Enhanced inhibitory control should result in less pronounced N200 amplitudes in simultaneous interpreters compared to controls. At the point of this dissertation, only two studies, one by Dong & Zhong (2017) and one by Zhao et al. (2024) have investigated the interpreter advantage hypothesis (see Chapter 6.3) in context of the Flanker by using electrophysiological measures. Dong & Zhong (2017) compared interpreters with more interpreting experience to interpreters with less interpreting experience in a Flanker task. On top of a smaller interference effect in reaction times, they found larger N200 amplitudes in both congruency conditions for interpreters with more interpreting experience and smaller P300 amplitudes in the incongruent condition. They interpreted this as evidence of monitoring and inhibition advantages for interpreters with more interpreting training (Dong & Zhong 2017: 195–202). It is debatable whether their findings of an increased N200 amplitude for the interpreters with more interpreting experience really reflects better inhibitory control or increased inhibitory control effort or more resource allocation towards inhibition. Although some studies provided electrophysiological evidence that there is a difference in executive processing between monolinguals and bilinguals, the interpretation of larger N200 amplitudes for bilinguals compared to monolinguals is not clear cut. Cespón (2021) emphasizes that the modulations of the N200 amplitude are frequently misinterpreted and that higher N200 amplitudes in bilinguals should not be automatically associated with better executive functions. As further above in this chapter, research such as by Jodo & Kayama (1992) has linked increased N200 amplitudes to greater cognitive effort and increased engagement of executive functions. So, while it is possible that bilinguals and interpreters use their executive functions more due to the increased processing demands of managing multiple languages, this does not automatically indicate a more efficient use. As reported in Chapter 6.3, Zhao et al. (2024) found less negative N200 responses for consecutive interpreters along with a lower error rate for interpreters compared to non-interpreting controls. Furthermore, dividing the Flanker task into two phases for training, they showed that interpreters also showed less negative N200 amplitudes in the second phase of Flanker trials compared to non-interpreters, another indication of that the interpreters allocated fewer cognitive resources to inhibitory control while also outperforming non-interpreter in terms of error rate but not in reaction time.

In the context of my empirical study it is therefore important to reason according to aforementioned research (see e.g. Zhao et al. (2024), Jodo & Kayama (1992), Cespón & Carreiras (2020), Markiewicz et al. (2023)). Therefore, a more negative N200 amplitude will be interpreted as increased cognitive resource allocation to inhibitory control (interference suppression) based on the aforementioned studies and a reduction of said amplitude across the course of testing while behavioural performance increases will indicate less cognitive resource allocation to inhibitory control and therefore less need to engage inhibitory control due to more efficient processing.

The empirical part of my dissertation extends the aforementioned electrophysiological research on conflict monitoring and inhibition with the Flanker task.

Although most studies report a peak on the N2 between 200-300ms after stimulus presentation, the time window of the N2 as well as its peaks vary from study to study.

In order to back up the decisions of the following statistical analysis of the Flanker task, Table 2 shows a selection of relevant studies that were either chosen because they used the Flanker task and investigated the N200 component in the context of the bilingual advantage or interpreter advantage or because they used a similar methodology applying the Flanker task in association with the N200 component. Table 2 presents N200 peak time windows as well as the analysed electrodes. Although all studies use variations of frontal and fronto-central electrodes, the common denominator are the central and fronto-central midline electrodes fz, fcz and cz that were chosen for areas of interest after visual inspection of single-electrode ERPs. Based on the reviewed time-windows and the general definition of the N200 peak time-window, a broad time window of 200-400ms was chosen for the Flanker task.

Table 2: Selection of studies using N200 in the Flanker tasks

Study	**N200 Peak Time window**	**Analysed electrodes**
Dong & Zhong (2017)	240-380ms	FPZ, AF4, AF3, F7, F5, F3, F1, Fz, F2, F4, FC5, FC3, FC1, FCz, FC2, FC4
Zhao et al. (2024)	250-390ms	F5, F3, FC5, FC3, F1, FZ, F2, FCz, F4, F6, FC4, FC6
Kousaie & Phillips (2012)	260-420ms	Fz, Fcz
Kousaie & Phillips (2017)	220-420ms	Fz, Fcz
Markiewicz et al. (2023)	275-325ms 325-375	Fcz as example
Purmann et al. (2011)	200-300ms	Fz, FCz, Cz, CPz
Heil et al. (2000)	155-370ms	Fz, Cz, Pz, C3, C4
Wild-Wall et al. (2008)	100-300	Fz, Fcz; Cz
Grundy et al. (2017b)	230-330ms	AF3, AFz, AF4, F1, Fz, F2, FC1, FCz, FC2

In addition to the research on N200 in conflict monitoring tasks such as go/no-go tasks and Flanker tasks, the N200 has also been associated with attentional switching in task and language switching paradigms and the concomitant suppression of interference. In an early study of task switching processes, Rushworth et al. (2002) investigated whether there are different cognitive processes at work during task switching. They found an N200-like modulation at 200-360ms after stimulus onset and a late positive modulation at 400-920ms after stimulus-onset. The authors concluded that there is more than one cognitive process involved in task switching, some rather linked to initiating and reconfiguring the task set before the performance of the task starts and some processes rather linked to the performance connected to the task item. This was also mentioned by Monsell (2003) who explained that switching from one task to another in switch trials requires task-set reconfiguration that involves switching attention between stimulus representations and task rules. He also highlights that this requires the inhibition of elements from the previous stimulus and the associated task rule and sees the task-set reconfiguration as a source for the switch costs (Monsell 2003: 135).

The processes initiating a new task and reacting to a stimulus are reflected in the cues and the target items of the task (Rushworth et al. 2002: 1139–1146). This differentiation has been summarized by Jamadar et al. (2015): As explained in Chapter 4.2, task switching paradigms usually differentiate between proactive control and reactive control. In the context of ERPs, different approaches can be mapped onto these two kinds of control. Proactive control is normally investigated by examining cue-locked ERPs, since they reflect the preparation interval that precedes a possible task switch. However, in this dissertation I focus on reactive control, that is investigated by examining so called target- or stimulus-locked ERPs which reflect processes involved in suppressing interference from previous trials, switching attention as well as conflict monitoring. These executive processes are often reflected in an N200 component with switch trials showing larger and sometimes delayed fronto-central N200 components compared to repeat trials (Jamadar et al. 2015: 331–333). As in the aforementioned modulations of the Flanker task, an increased N200 amplitude in task switching reflects increased effort and more resource allocation to inhibitory control (interference suppression) during attentional switching (Cespón & Carreiras 2020: 317).

Although not as frequently as reaction time measures, the N200 ERP component has also been used in task switching paradigms to investigate executive functions in bilinguals, often in relation to the bilingual advantage hypothesis or to show overlapping language and executive processing. López Zunini et al. (2019) for example investigated task switching abilities in younger and older bilinguals. Although they did not find a difference in effect size in cue-locked mixing and switch positivity or target-locked mixing and switch N200, they found that bilinguals showed larger target-locked N200 amplitude compared to monolinguals in both conditions, which they suspected to reflect increased conflict monitoring for bilinguals in all tasks. They also found larger N200 amplitudes for older bilinguals compared to younger bilinguals. They suggested that bilinguals rely on different processing strategies than monolinguals as they age. In terms of the behavioural data, they found smaller mixing and switch costs in bilinguals, which they interpreted towards a better ability to focus attention and reconfigure task specific processes for bilinguals than monolinguals (López Zunini et al. 2019: 8–9).

Timmer et al. (2017) examined whether there are differences in processing in task- and language switching between monolinguals and bilinguals. They found that monolinguals and bilinguals were similar in their behavioural switching and mixing costs in task switching, but that bilinguals outperformed monolinguals in all conditions of the mixed blocks. In terms of ERPs, more distributed processes could be shown for bilinguals than for monolinguals in both language and task switching. The ERP components N200 and P300 were more similar in bilinguals than in monolinguals which was interpreted as an increased processing overlap for bilinguals. The authors connected this to the sustained training in language switching that is supposed to alter switching processes in verbal as well as domain-general domains (Timmer et al. 2017: 254).

As reviewed in Chapter 6.3, there are studies that investigated task-switching in simultaneous interpreters with behavioural data. However, electrophysiological research on domain-general task switching that specifically examine the N200 component in simultaneous interpreters is lacking.

Table 3 shows a selection of ERP studies that were chosen since they either investigated the N200 component in association with attentional switching by applying a task switching paradigm in the context of the bilingual advantage hypothesis or because they used a similar methodology to the one chosen in this dissertation. Table 3 presents the corresponding N200 peak time windows and the analysed electrodes. In addition to central and fronto-central electrodes, these studies tended to also use parietal as well as temporal electrodes. The central and fronto-central electrodes however were again a common denominator across studies and, after visual inspection of the single-electrode ERPs in this study, they were again chosen as most reliable. In comparison to the Flanker task, studies using switching paradigms seemed to use more narrow time windows to investigate the N200, therefore and after visual inspection of the data specific to this study, a time-window of 200-320 ms after stimulus presentation was chosen.

Table 3: Studies that used N200 in switching tasks

Study	**N200 Peak Time window**	**Analysed electrodes**
Rushworth et al. (2002)	200-360ms	Fpz, Fz, Cz, Pz, Oz, Fc1, Fc2, O1, O2, FP1FP/2, AF3/AF4, AF7/AF8, F5/F6, F7/F8, FC1/FC2, FC3/FC4, C3/C4, CP1/CP2, CP3/CP4, FC5/FC6,FT7/FT8,T7/T8,CP5/CP6, TP7/TP8, P5/P6, P7/P8, PO3/PO4, PO7/PO8, O1/O2
Timmer et al. (2017)	275-300ms 225-275ms	C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2, PO3, POz, PO4
López Zunini et al. (2019)	200-400ms	Fz, Fcz, Cz, F3, Fc3, C3, F4, Fc4, C4, Cpz, Pz, Oz, CP3, P3, O1, O2, C4, Cp4
Declerck et al. (2021)	200-350ms	Fpz, Fp1, Fp2, F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, O2

The P300 component was first mentioned by Sutton et al. (1965) as a late positive component that peaks around 300ms in the context of certainty and uncertainty in stimulus presentation (Sutton et al. 1965: 1187). In general the P300 has been considered to reflect stimulus categorization (Bialystok 2017: 245) or updating memory after a stimulus was categorized (Daffner et al. 2011: 1299). P300 has been associated with inhibition processes in stimulus-response compatibility (SRC) tasks such as go/no-go tasks or Flanker tasks as well as the allocation of attentional resources during attentional switching and working memory processes (Cespón & Carreiras 2020: 316). As explained in a review of the P300 by Polich (2007), stimuli that have to be kept in memory during a working memory task can yield larger P300 for recurring stimuli than for new stimuli. Furthermore, the size of the P300 amplitude is a product of the extent to which attentional resources are allocated during dual-task paradigms. This is based on the model for the allocation of attentional resources by Kahneman (1973) that was already briefly mentioned in Chapter 6. In this model, an overall arousal level determines how much attention is available to perform a task. Tasks that need more attentional resources are reflected in smaller P300 peak amplitudes and longer peak latencies (Polich 2007: 2129–2131). Polich (2007) summarizes that P300 peak latencies indicate how fast a stimulus can be classified and are modulated by task processing demands. He also related the P300 latency to working memory updating processes (Polich 2007: 2129–2132). Cespón & Carreiras (2020) explain that SRC-tasks entail switching attention from one stimulus-response mapping to another. Longer P300 latencies have been reported for trials that required switching from the previous stimulus-response mapping to another than in trials that require only the retrieval or repetition of the previous stimulus-response mapping (Cespón & Carreiras 2020: 318).

The P300 has often been divided into two sub-components: The P3a that is also called novelty P3 and is rather fronto-centrally distributed, and the P3b that is rather temporal-parietally distributed (Polich 2007: 2141). The fronto-central P300 has been associated with response inhibition processes in Go/No-Go tasks. More specifically, larger fronto-central P300 amplitudes have been associated with more difficulty to inhibit certain type of stimuli, usually conflict trials relative to non-conflict trials. Cespón & Carreiras (2020) summarize that the P300 amplitude is generally larger in non-conflict trials (such as repeat trials in the case of task-switching) relative to conflict stimuli trials (such as switch trials in task-switching paradigms). The P300 has also been associated with attentional switching, with larger parietal P300 amplitudes for repeat than for switch trials and larger parietal P300 amplitudes for lower task difficulty in switching tasks (Cespón & Carreiras 2020: 317–318). When a switch trial occurs in a task-switching paradigm, attentional resources must be allocated to switch to updated stimulus-response mapping.

In task switching paradigms, target-locked ERPs show a smaller centroparietal P300 in switch trials compared to repeat trials in addition to the larger fronto-central N200 mentioned in the previous chapter. While the N200 amplitude grows as interference level grows, the P3b amplitude is decreased (Jamadar et al. 2015: 333). In task switching and SRC paradigms, the N200 and P3b are frequently so closely linked that literature often talks about an N2/P3 complex (Karayanidis & Jamadar 2014: 219). In SRC tasks, such as the Flanker task, P300 reduces while attentional load is increased (Jamadar et al. 2015: 333).

Therefore, when task switching is applied in bilingualism literature, P300 is frequently reported in combination with N200, such as in the earlier mentioned study by López Zunini et al. (2019). In their study, bilingual adults showed smaller P3b amplitudes than monolingual adults in mixing costs. The authors assumed that ageing bilinguals rely stronger on earlier processes and have automated some processing strategies compared to monolinguals (López Zunini et al. 2019: 7).

In the aforementioned study by Timmer et al. (2017), P300 modulation was shown for both monolinguals and bilinguals. However, contrary to literature, they found larger positive P300 amplitudes for switch than for repeat trials. They note that P300 is present in parietal and parietal-occipital regions but observe an extension to central-parietal regions for bilinguals only. Those P300 regions were active during task switching and, for bilinguals, more of those regions were also active during language switching, inferring a larger amount of cross-domain overlap for bilinguals (Timmer et al. 2017: 255). In terms of latency, Kousaie & Phillips (2012) found later P300 peak latency for monolinguals compared to bilinguals in a Stroop task but this was not specific to congruent or incongruent trials (Kousaie & Phillips 2012: 76). Literature in studies investigating the P300 in the context of switching in interpreters and translators is lacking.

Table 4 shows a selection of ERP studies that were chosen since they investigate attentional switching in the context of task switching paradigms (mostly colour-shape switch) with the P300 component in the context of the bilingual advantage hypothesis or because they were comparable in their methodology to the design applied in this dissertation. Table 4 presents the respective P300 peak time window and the analysed electrodes. In accordance with relevant P300 literature (see Cespón & Carreiras (2020), Sutton et al. (1965), Polich (2007)), these studies use mostly frontal, central and parietal electrodes and a time-window between 300 and 600ms after stimulus presentation. Accordingly, after visual inspection of the single-electrode ERP data in this study, the electrodes cz, cp1, cpz and p1 were used as an area of interest for analysis in a time window of 280-400ms after stimulus onset.

Table 4: Studies using P300 in task switching

Study	P3 peak window	Analysed electrodes
Periáñez & Barceló (2009)	310-730ms	Fz, Cz, Pz
Timmer et al. (2017)	325-400ms 300-375ms	C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2, PO3, POz, PO4
López Zunini et al. (2019)	300-600ms	Fz, Fcz, Cz, F3, Fc3, C3, F4, Fc4, C4, Cpz, Pz, Oz, CP3, P3, O1, O2, C4, Cp4
Declerck et al. (2021)	400-600ms	Fpz, Fp1, Fp2, F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, O2
Chen et al. (2022)	330-500ms	F1, Fz, F2, C1, Cz, C2, P1, Pz, and P2

The parietal P300 has also been associated with updating working memory (see Polich (2007); Cespón & Carreiras (2020)). Polich (2007) summarizes that upon seeing a stimulus, a comparison process that is lead by attention occurs to compare the previous stimulus in working memory to the current stimulus. If the new current stimulus does not match the previous one, attentional resources must be allocated to update the stimulus representation (categorize the stimulus) in working memory which is shown in the P300 amplitude (Polich 2007: 2129).

The n-back task (Smith & Jonides 1997) is a common task to successfully measure working memory updating processes. The n-back task is thought to represent the execution of different operation sets. One operation set concerns maintaining and manipulating contents in working memory while the other set of operations involves determining whether the presented stimulus is a match based on whether it corresponds to a representation that is still retained in working memory (Daffner et al. 2011: 1299). Kok (2001) discusses that when stimulus categorization becomes more difficult, the P300 latency increases (Kok 2001: 557) and assumes that as task difficulty increases, the P300 amplitude decreases (Kok 2001: 571). As reviewed in Cespón & Carreiras (2020), the amplitude of the parietal P300 is larger with decreasing working memory load or task difficulty in working memory tasks (Cespón & Carreiras 2020: 318). Similar results were found specifically for the n-back task in Daffner et al. (2011) who postulated that with increasing task demands in the n-back task, the P300 would decline since more cognitive resources need to be allocated to sustain attention and manipulate information than for decision-making and updating. They tested this hypothesis by comparing high and low performing younger and older adults on 3 levels of the n-back task. They found that especially for low performers, increasing task demands resulted in a decrease in the P300 amplitude (Daffner et al. 2011: 1299–1308).

Studies that investigate the P300 component in the context of working memory as an executive function in bilinguals are not as numerous as studies on other executive functions. However, studies such as by Morrison et al. (2019) and Morrison & Taler (2020) provided evidence of how bilingualism modulates the P300 response. As already described in Chapter 5.4 in the context of the bilingual advantage, Morrison et al. (2019) found larger P300 amplitudes for bilinguals compared to monolinguals in the context of an n-back task, indicating that the task was less difficult for bilinguals. However, there was no significant group difference in reaction times and accuracy. Morrison & Taler (2020) used ERP to compare the performance of young and older monolinguals and bilinguals on a delayed matching-to-sample task. In conditions of low, medium and high working memory load, participants saw a maximum of five number on a computer screen and were shown another number after a delay period. They then had to indicate whether the number shown after the delay period was represented on the screen beforehand. For the N200 component, results showed that the N2 amplitude increased as working memory load increased and that monolinguals showed larger N2 amplitudes than bilinguals in all working memory load conditions. Furthermore, young adults displayed larger N2 amplitudes than older adults. The study differentiated between fronto-central P300 and parietal P300. For the fronto-central P300 component, the P3 amplitude was significantly smaller in the high load condition than in the medium and low load condition. Additionally, monolinguals showed smaller P300 amplitudes than bilinguals. Finally, older adults had larger P300 amplitudes than younger adults, especially in the medium and high load condition. For the parietal P300 component, the P3 amplitude also significantly decreased as working memory load increased. However, there were no significant differences due to language or age overall. Results indicated however, that the P3 amplitude was larger in young adults compared to older adults in the low load condition. In terms of the behavioural results, there was no significant difference between monolinguals and bilinguals in reaction times and accuracy and there was no interaction of age and language. However, older adults showed significantly longer reaction times in general, accuracy did not differ in terms of age. Morrison & Taler (2020) interpreted the N200 component as a reflection of increased effort when a task is difficult. Since monolinguals showed larger N200 amplitudes than bilinguals, they proposed that monolinguals needed more effort to perform the task compared to bilinguals. However, the result that older adults showed smaller N2 amplitudes than younger adults and therefore seemed to perform the task with less effort was not consistent with the theoretical background of the aging brain. The P300 component was interpreted as a reflection of available resources and working memory, with larger amplitudes associated with more available resources to perform the task. Both parietal and anterior P300 amplitudes decreased with increasing working memory load, however, only the frontal P300 amplitude was larger for bilinguals than for monolinguals and larger in older adults than in young adults. It therefore seems that bilinguals had more resources available to solve the working memory task. Nevertheless, the interpretation of less effort and more available resources for bilinguals over monolinguals was not reflected in the performance indicated by reaction times and accuracy (Morrison & Taler 2020: 2–8).

Dong et al. (2015) compared monolingual participants with different working memory capacities categorized by a digit span task in their performance of an n-back task with different difficulty levels. They found that the P300 amplitudes decreased with increasing working memory load and larger amplitudes at parietal sites (Dong et al. 2015: 148–150).

Barker & Bialystok (2019) compared voltage in the P300 time-window between monolinguals and bilinguals performing an emotion 2-back and a 3-back task. They found an effect of memory load with the P300 amplitude in the 2-back task being smaller than the amplitude of the 1-back task for all participants. Additionally, they found longer latencies on match trials than on no-match trials in the 1-back task but not in the 2-back task. Unfortunately, they did not find a main effect of language group but groups differed on match-trials in the 2-back task (Barker & Bialystok 2019: 31–37).

Similarly, Comishen & Bialystok (2021) compared a monolingual and a bilingual group in an n-back task with task demands increasing from 0-back to 3-back. They found a main effect of group with the bilingual group showing larger overall P300 amplitudes than the monolinguals. Furthermore, they found a main effect of condition with a decrease of the P300 amplitude in the 2-back task and a further decrease of amplitude for the 3-back task. Additionally increased task performance was associated with larger P3 amplitudes. Based on these results, the authors suggest that bilingualism increases the efficient use of attentional resources (Comishen & Bialystok 2021: 3–8).

To the extent of my knowledge, studies using the P300 ERP component to study working memory updating in interpreters or translators have yet to be conducted.

Table 5 shows a selection of ERP studies that were chosen since they use the P300 in association with working memory updating in the context of the bilingual advantage hypothesis or because they applied a similar methodology to the one applied in this dissertation. Table 5 summarizes their P300 peak time windows and the analysed electrodes. Again, studies used both central, frontal and also parietal electrodes oriented on relevant P300 literature (see Cespón & Carreiras (2020), Sutton et al. (1965), Polich (2007)) and applied time-windows spanning from 280-500ms after stimulus presentation. Based on this review and visual inspection of the single-electrode ERPs, an area of interest consisting of the electrodes cp1, cpz, p1, pz and poz was summarized and voltage was analysed in a time-window between 280-400ms in this dissertation.

Table 5: Studies using P300 in working memory tasks

Study	**P300 peak window**	**Analysed electrode**
Dong et al. (2015)	280-500 ms	F3,FZ, F4, P3,PZ,P4
Morrison et al. (2019)	300-500 ms	Fz, FCz, Cz, and Pz
Barker & Bialystok (2019)	300–450 ms	p1, Cpz, Cp2, P1, Pz, P2
Morrison & Taler (2020)	250-450 ms	P3, Pz, P4
Comishen & Bialystok (2021)	325-425 ms	CP1, CPz, CP2, P1, Pz, P2, PO3, POz, PO4
Daffner et al. (2011)	275-50ms	Cz, pz

8.1.3 Participants

With the aim of investigating how learning to interpret affects different executive functions, two groups of participants were recruited from the MA Interpreting at the Fachbereich Translation, Sprache und Kultur (FTSK)[4] of the Johannes-Gutenberg University Mainz in Germersheim. Other studies such as by Morales et al. (2015), Yudes et al. (2011), Henrard & van Daele (2017) or van der Linden et al. (2018) assume that professional simultaneous interpreters must possess enhanced executive functions in comparison to non-interpreting populations and therefore compare interpreters to professional translators or language teachers in a cross-sectional design. However, this dissertation investigates whether and when executive functions are enhanced by SI training and therefore examines student interpreters through their process of learning how to interpret simultaneously in a longitudinal design.

Additionally, a control group of students enrolled in the MA Translation at the Fachbereich Translation, Sprache und Kultur (FTSK) of the Johannes-Gutenberg University Mainz in Germersheim was tested for baseline comparison to ensure that both groups possessed the same cognitive abilities at the beginning of their

---

[4] Translation Studies, Linguistics and Cultural Studies

respective training. The application process for the MA Conference Interpreting at FTSK requires an aptitude test, the likes of which was discussed in Chabasse (2009). It could be argued that by passing this test, students with already enhanced abilities that have the potential to become good interpreters are chosen. Therefore, it is important to compare the interpreting group with a control group (in this case the translators) at baseline to see if the interpreting group shows an inherent advantage in executive functions (see also Babcock et al. (2017)). Translation students were chosen over other bilingual populations since studies such as by Prior & Gollan (2011) and Verreyt et al. (2016) have shown that how often bilinguals switch between languages seems to be an indicator of whether executive functions are affected or not. The decisive difference between written translation and simultaneous interpreting lies in the urgency that is created by the near simultaneity of language reception, translation and production processes during simultaneous interpreting (see e.g. Frauenfelder & Schriefers (1997)). These overlapping processes are assumed to put much strain on executive processing that is involved in managing the different languages (see Hervais-Adelman & Babcock (2020)). However, translation students also switch between languages daily during translation, more than bilinguals without professional training who speak multiple languages but do not engage in language switching as intensively as during SI or written translation. Apart from their differential language switching habits, all participants fell under the category of late sequential bilinguals for the language combination German as first language and English as second language.

The Interpreting group initially consisted of 10 participants and the Translation group consisted of 12 participants. This participant number is quite low, which was due to the complexity of the experimental manipulation; participants had to commit to a two-year EEG study and only a limited number of participants qualified to participate in this study according to their language combination. However, participant numbers this low are not entirely uncommon in studies that use neuroscientific measures such as EEG and fMRI, such as studies by Elmer et al. (2010) that had 11 participants per group for their EEG-study, Elmer & Kühnis (2016) who had 12 participants per group and Morales et al. (2015), who had 16 participants per group for their respective fMRI studies.

Before participating in the study, all participants were required to fill out the Translation and Interpreting Competence Questionnaire (TICQ) (Schaeffer et al. 2020). The TICQ is a multilingual tool that is available online, can be customized and is aimed at collecting quantitative and qualitative data on different aspects of linguistics skills, translation and interpreting skills. Contrary to other measures of translation and interpreting competence it is not based on non-validated ad hoc instruments but was validated in two separate studies. It aims at furthering comparability across studies, an issue pertaining to both bilingualism studies as well as translation and interpreting studies (see Chapter 3 and Chapter 6.1). Consisting of three sections, Section A covers general participant information such as age, gender, and course of studies as well as aspects pertaining to language history and proficiency such as Age of Acquisition, years of learning, different modalities of using languages as well as self-rated proficiencies. Section B is translation specific, and Section C is interpreting specific. For the study in this dissertation, Section C was used. It includes questions about the participants professional experience, interpreting competence and questions regarding the interpreting process such as strategies and research (Schaeffer et al. 2020: 91–96).

The TICQ was used to match participants in both groups as well as to compare proficiency in the interpreting group at baseline and at the end of their 4 semester training period. The TICQ results from the training period of the interpreting group are presented in Chapter 9.2.1.

For baseline comparison of both groups, participants from both groups were matched on age, gender, age of English acquisition, self-rated English proficiency and self-rated interpreting proficiency. Two-sample t-tests indicated that there are no group differences in all aforementioned variables at baseline, which is a prerequisite for interpreting the behavioural as well as ERP data at baseline between groups. Table 6 gives an overview over this participant data.

Table 6: Between-group matching for age, gender, AoA English, English proficiency, interpreting proficiency

	Interpreting group	**Translator group**	**Group difference**
Age	Mean: 24, SD: 2.5495	Mean: 23.92, SD: 3.6296	t = 0.058702, df = 19, p-value = 0.9538
Gender	Mean: 1.0, SD: 0.0	Mean: 1.1667, SD: 0.3892495	t = 1.4832, df = 11, p-value = 0.1661
Age of EN acquisition	Mean: 8.33, SD: 2.0615	Mean: 8.08333, SD: 2.7784	t = 0.22662, df = 19, p-value = 0.8231
Self-rated EN proficiency	Mean: 76.11, SD: 12.444	Mean: 75.83333, SD: 10.62444	t = 0.053845, df = 15.721, p-value = 0.9577
Self-rated interpreting proficiency from L1 into L2	Mean: 44.4444, SD : 19.59663	Mean: 37.25, SD: 32.30008	t = 0.63194, df = 18.365, p-value = 0.5352
Self-rated interpreting proficiency from L2 into L1	Mean: 61.6667, SD: 18.20027	Mean: 44.75, SD: 34.83239	t = 1.4405, df = 17.312, p-value = 0.1676

Of all participants, only one participant from the control group indicated to have been fluent in English before the age of seven. All participants indicated to have acquired English in the context of formal education and furthermore, all participants indicated that they spoke English between 1 and 15 hours a week.

All participants from the Control group (the Translator group) indicated that they were currently enrolled in the MA Translation at FTSK Germersheim. All participants in the Experimental group (the Interpreting group) indicated that they were currently enrolled in the MA Conference Interpreting at FTSK Germersheim. Participants from both groups indicated that they had not had any prolonged experience in simultaneous interpreting immediately before starting their respective MA programmes. Some participants from both groups indicated that they had participated in trial courses of simultaneous interpreting or tested SI within a single semester. Furthermore, participants from both groups indicated that they had once interpreted for a family member, which was also not rated as extensive and continuous simultaneous interpreting training. This information reflects how difficult it is to find a group that has not yet had any contact to any interpreting action, since especially the FTSK is an environment that fosters multilingualism as equally as applying multilingual and translation skills. Most importantly, none of the participants from both groups had gone through several semesters of simultaneous interpreting training at baseline.

Only the interpreting group was tested longitudinally across four semesters, the designated duration of the MA Conference Interpreting at the FTSK Germersheim. Originally, I planned to test both the Experimental group as well as the Control group longitudinally, however, due to Covid-19 restrictions and ensuing restricted lab use, only the interpreting group could be tested longitudinally. According to the module handbook of the Master of Conference Interpreting at FTSK Germersheim (Fachbereich 06 2021), participants must take a set of mandatory modules that consist of a module for general Interpreting Studies, a module for Cultural Studies, a module for Linguistics and finally six modules containing four applied classes for consecutive and simultaneous interpreting each. The workload for each of these applied interpreting modules is listed as 360 hours. Simultaneous interpreting makes up 50% of a module, consecutive interpreting the other 50% and both interpreting modes are listed with the same workload. It is therefore evident from the module handbook that all participants in this study received the same amount of simultaneous interpreting training which amounts to a workload of 180 hours per module, so 1080 hours in total for the whole Master of Conference Interpreting.

For the interpreting group, another prerequisite was that their mother tongue needed to be German, and English was supposed to be their first or second language of interpreting (they received equal amounts of interpreting training in their first and second working language). Participants needed to indicate how many hours per week they spend time interpreting, their ratio of simultaneous vs. consecutive interpreting, their maximum interpreting capacity per day as well as some questions about the interpreting process itself concerning preparation, problem solving strategies, estimated décalage or use of resources in the booth.

Since the MA interpreting programme inherently has a limited number of first-semester students and the conditions for participation in the study were also very specific, only ten participants could be recruited for the Experimental group.

All participants' mother tongue was German and all had grown up in Germany. All participants in the experimental group indicated English as their L2, expect for one participant who indicated French as L2 and English as L3. All participants are late bilinguals, meaning that none of them was fluent in their L2 or L3 before the age of seven. All participants learned their L2 in the context of formal education. Participants have been learning their L2 for an average of 14.1 years (SD = 2.38).

Further analysis of participant information of the interpreting group that is related to interpreting competence is reported in Section 8.2.1, where the competence change from baseline to the end of the training programme is examined.

8.1.4 Stimuli

The Flanker task stimuli consisted of 5 arrowheads that were horizontally arranged. The middle arrowhead was flanked by two arrowheads, either pointing in the same direction as the middle arrowhead or in a different direction. This scheme created two congruent conditions (<<<<<, >>>>>) and two incongruent conditions (<<><<, >><>>). The Flanker experiment design was based on the experiment by (Dong & Zhong 2017). There were five blocks of 72 trials each. The congruent/incongruent ratio in each block was 75%/25%. This ratio was chosen based on studies by Purmann et al. (2011) and Costa et al. (2009). Purmann et al. (2011) examined adaption to frequent conflict in a Flanker task by varying frequency of incompatible stimuli. They found that the fronto-central N2 amplitude was reduced under frequent conflict as opposed to less frequent conflict (Purmann et al. 2011: 56). Costa et al. (2009) investigated the bilingual advantage in the Flanker task under different monitoring effort conditions. In high monitoring contexts (e.g. 75% congruent/25% incongruent trials), bilinguals outperformed monolinguals in terms of reaction times but not in low monitoring versions (e.g. 92% congruent/ 8% incongruent) (Costa et al. 2009: 141).

Participants were told to only focus on the middle arrowhead. Using both hands, participants pressed the "F" key if the arrowhead pointed to the left side and the "J" key if the arrowhead pointed to the right. Before each trial, a fixation cross appeared on the screen for 350-550ms. This time interval was jittered to prevent the participants from building up a "habituation effect" that occurs when stimuli are repeated continuously and in a monotone manner, which causes the electro-negative components to be attenuated. However, the amplitude is restored if two stimuli are associated (Walter et al. 1964: 380). The fixation cross acts as a so-called *indicative stimulus* that operates as a warning for the following *imperative stimulus*, to which the participant is often required to make a response. By jittering the fixation cross time, participants were prevented from building up expectancy. The fixation cross was followed by a blank space that lasted 300ms. After that, the stimuli appeared on screen for 750ms or until the participant had pressed a key to answer. The participant had to answer within the 750ms. After the stimulus had disappeared, there was a jittered interstimulus interval for 900-1100ms (Walter et al. 1964). The experimental blocks were preceded by a practice block that contained 12 practice trials with a 50/50 congruent/incongruent ratio. A flowchart of the Flanker task is presented in Figure 1:

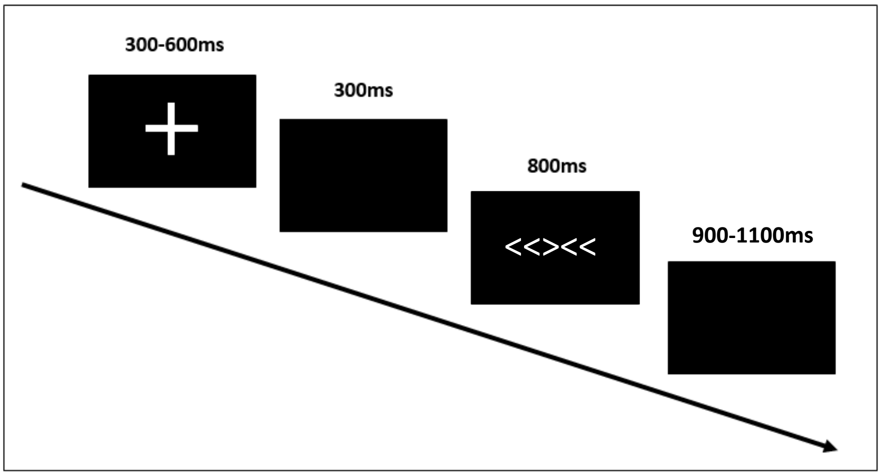


Figure 1: Flowchart Flanker task

Derived from the experimental designs in López Zunini et al. (2019) and Prior & Macwhinney (2010), stimuli for the colour-shape switch task consisted of circles and triangles in red or blue. Every stimulus was preceded by a cue that indicated whether the participant was supposed to react to the shape, or the colour of the stimulus presented. The colour cue was a colour gradient, and the shape cue were four small shapes that were presented horizontally. Every experiment consisted of two single task blocks of 56 trials each, in which the participants either reacted to the shape or the colour. The two single task blocks were then followed by four mixed task blocks consisting of 56 trials each in which the participants were supposed to react to the shape, or the colour of the stimulus as indicated by the preceding cue. Before a cue was presented, a fixation cross appeared in the centre of the screen for 300-350ms followed by a blank screen for 150ms seconds. After that a shape or a colour cue appeared for 250ms followed by the stimuli that remained on the screen for 4000ms or until the participant had pressed a key. Subsequently, an intertrial interval followed for 850ms until the next fixation cross appeared. A flowchart of the single task trials can be seen in Figure 2 and a flowchart of the mixed task trials can be seen in Figure 3.

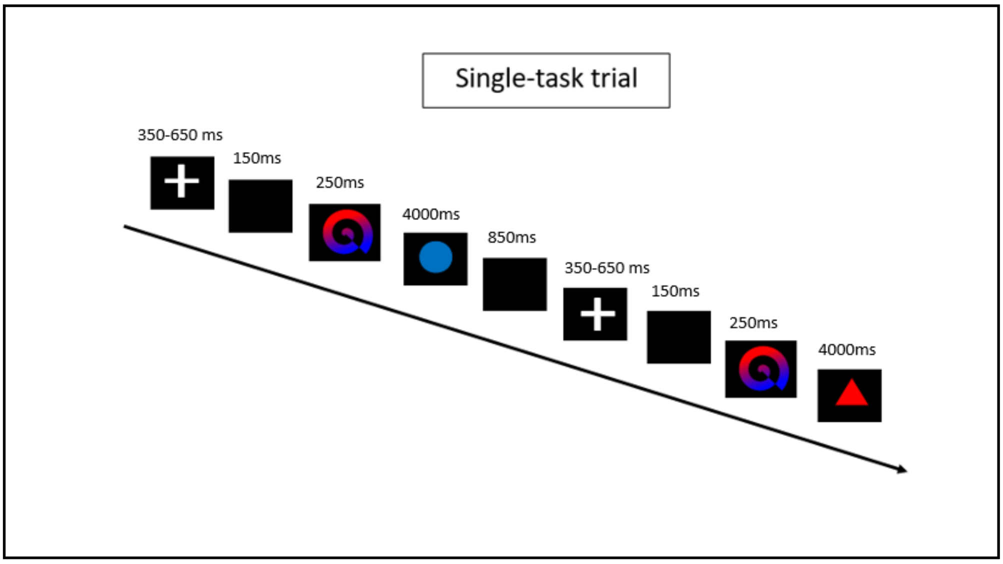


Figure 2: Flowchart single task trial

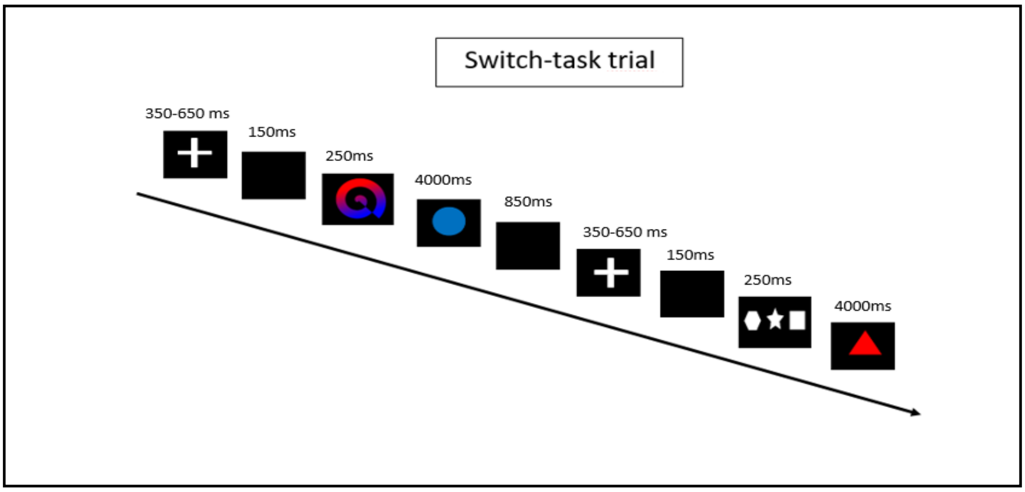


Figure 3: Flowchart mixed task trial

The participant used both hands to react to the stimuli. They used their left hand and consequently the "S" key to react to a circle and the "D" key to react to a triangle. They used their right hand and the "K" key to react to the colour "red" and the "L" key to react to the colour blue.

Based on the design of Szmalec et al. (2011), the stimuli in the 2-back and 3-back tasks consisted of 20 consonants of the alphabet (D, H, M, T, B, G, R, Z, P, S, F, J, K, V, N, W, L, X, C, Y). The materials were held in close comparison to the ones in Szmalec et al. (2011). In the current experiment, participants first completed a 2-back task followed by a 3-back task. In both versions, participants saw a string of letters that appeared in the centre of the screen and had to press "B" on the keyboard whenever a letter was the same as two letters (2-back) or three letters (3-back) before. Each stimulus appeared on screen for 500ms, followed by a jittered blank screen of 800-1000ms. A flowchart of the 2-back and 3-back task can be seen in Figure 4:

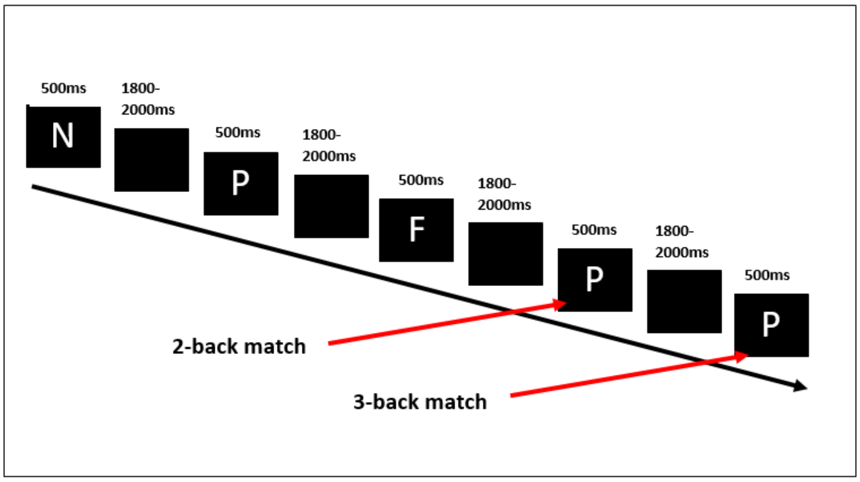
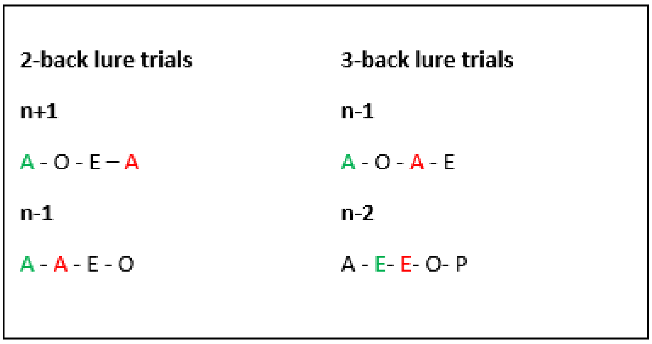


Figure 4: Flowchart 2-back and 3-back task

In total the n-back task consisted of 12 blocks (6 blocks 2-back and 6 blocks 3-back). The 2-back task had 47 stimuli, since letter matching could only start at the 3rd letter and the 3-back task had 48 stimuli, since letter matching could only start at the 4th letter. Every block contained 30 mismatch and 15 match trials. Of the 30 mismatch trials, 6 were lure trials. For the 2-back condition, 3 lure trials were in the n+1 position and 3 were in the n-1 position, while for the 3-back condition, 3 lure trials were in the n-1 position, and 3 lure trials were in the n-2 position. Examples of the construction of lure trials in both conditions can be seen in Figure 5:



2-back lure trials	**3-back lure trials**
n+1	**n-1**
A - O - E – A	A - O - A - E
n-1	**n-2**
A - A - E - O	A - E- E- O- P

Figure 5: Examples of lure trial construction for 2-back and 3-back conditions

For the analysis, the so called lure interference effect was examined. According to Szmalec et al. (2011), lure trials create competition between assessing how familiar an item as and how well it can be remembered. In the case of the 2-back task, an n+1 lure trial depicts the interference that is created by an old item that is not task relevant anymore and can be released from WM. On the contrary, an n-1 lure trial depicts the interference from an item that is still task relevant. The difference between the two lure trial types is that n-1 lure trials still have to be kept activated in WM as the item still might need to be processed while n+1 lure item is allowed to fade from WM. In the case if the 3-back task, an n-1 lure basically corresponds to a 2-back task that is however not required in the 3-back condition. The n-2 lure items correspond to 1-back tasks. According to Szmalec et al. 2011, both n-1 and n-2 lure trials require the participant to maintain the item active in the WM, still awaiting a possible match. In this waiting stage, participants rehearse those items in their mind. Items in the n-1 position in 3-back tasks are slightly older that items in the n-2 position and have therefore been rehearsed more often. This rehearsal to maintain an item in WM increases the familiarity of it. Szmalec et al. (2011) therefore predict that a lure trial in the n-1 condition creates more interference than a lure trial in the n-2 position, since it has been rehearsed more in WM and is therefore more familiar (Szmalec et al. 2011:7-9).

8.1.5 Data collection and apparatus

Before the study was launched, it was piloted by testing the EEG-setup, the equipment and the smooth display of the computer-based tasks within a trial run in the Neurolab at FTSK Germersheim.

Before the start of the first round of data collection, every participant was informed about the procedure of the experiment, the use of the EEG and data protection. After the participants had given informed consent, the EEG cap was fitted. Participants were seated in front of a computer screen and a keyboard in an EEG chamber that fulfilled the requirements of a Faraday cage to reduce noise. The experiment was coordinated and supervised by me from the outside of the EEG chamber on two separate screens. One screen mirrored the screen in the EEG-chamber and one screen showed the continuous EEG recording. The EEG was continuously recorded with an electrode cap with 64 Ag/AgCl electrodes that are placed conforming to the international 10-20 system. The SynAmps 2 system of Neuroscan (Compumedics Neuroscan) was used at a sampling rate of 1000Hz. Furthermore, one electrode was placed at each mastoid and the EEG signal was later re-referenced to the average. Impedances were kept below 5 kOhm at all electrode sites before and during recording.

After recording, the raw EEG data was processed using EEGLAB (Delorme & Makeig 2004) in the MATLAB environment (The MathWorks Inc. (2022)). A high-pass filter at 0.5 Hz and a low-pass filter at 40Hz was used for filtering. The raw EEG data was first cleaned manually by cutting off long pauses containing artefacts at the beginning and at the end of recording as well as between the blocks. Noisy channels were rejected manually, and independent components were computed using independent component analysis (ICA). Consequently, ICA-components that contained artefacts were rejected as well. Interpolation was used to replace the channels that were initially rejected. The data was divided into epochs which were cleaned again by manual rejection of noisy epochs. Finally, epochs were sorted according to the conditions of the experimental task, e.g. congruent and incongruent in the Flanker task. Only epochs in which the participants answered correctly were included. In order to analyse the ERP, epochs were defined from 100ms before stimulus onset to 700ms after stimulus onset.

8.1.6 Data Analysis

Data analysis covers three main parts: analysis of competence data, behavioural analysis of reaction times and error rates, and analysis of EEG data.

Competence data is initially analysed by comparing participant's answers to the pre- and post-training TICQ. Additionally, individual interpreting competence scores were determined for each participant: In the TICQ, certain questions regarding the translation or interpreting competence were assigned weights. It is possible to derive a summed competence score per participant that consists of weighted values of the answers to the questionnaire. During its validation, the robustness of the section to discriminate between lay bilinguals, students and professionals was tested by using multigroup discriminant function analysis (MDA). MDA revealed two discriminant functions, called F1 and F2, that were shown to both discriminate successfully between groups and predict to which group participants belonged to (e.g. translators, interpreters, bilinguals). The summed competence values per participant represent the F1 and F2 values (Schaeffer et al. 2020: 97–100). In Jacob et al. (2024), increased F1 values have been associated with increased self-rated competence values (Jacob et al. 2024: 552). While individual skill associated with the translation process is reflected by F1, F2 is associated with the skills of quality control and product evaluation (Jacob et al. 2024: 548–549).

Behavioural data in the form of reaction times and accuracy scores is analysed by using descriptive statistics as well as linear mixed regression models.

EEG data was analysed using the event-related potential (ERP) method. This includes comparing voltage in different conditions and sessions in the respective ERP-time windows by means of both descriptive statistics as well as linear mixed effects models and topographic ANOVAs. Furthermore, the EEG data analysis includes peak latency analysis as well as frequency analysis from a single participant. All linear mixed regression models were fitted with the lme4 package Bates et al. (2015) in the R environment.

Collinearity can have an adverse effect on the results of the linear mixed regression model. If collinearity is present, mostly if there are highly correlated predictors, one predictor can be predicted by other predictors which impairs the interpretation of the regression model (Winter 2020: 112). To verify the absence of multicollinearity, a variance inflation factor (VIF) was calculated. The VIF measures to what degree the independent variables are responsible for collinearity (Winter 2020: 114). Following Zuur et al. (2010), VIF-values that are equal to 1 indicate that there is no collinearity while values over 3 indicate collinearity or even strong collinearity the higher the value becomes (Zuur et al. 2010: 9).

In order to determine effect sizes, the Cohen's f^2 was calculated. Usually, Cohen's d is the go-to measure for calculating effect sizes. The Cohen's d determines the difference between two means that is divided by the standard deviation of the two groups, the latter representing the overall variability in the data (Winter 2020: 159). However, linear mixed models are more complex than basic approaches of significance testing and contain variances from various sources (e.g. random effects, fixed effects). The Cohen's d is therefore not suited for more complex data structures such as repeated observation of the same subjects as is the case in the study at hand (see Selya et al. 2012:1). Selya et al. (2012) propose the Cohen's f^2 to calculate the effect size in multiple regression models. Cohen (1977) set the guidelines to interpreting the Cohen's f^2 as a value of 0.02 as a small effect, a value of 0.15 as a medium-sized effect and a value of 0.35 or larger a as large effect (Cohen 1977: 413–414). After reviewing ERP literature that applies the relevant tasks (see Section 8.1.2 for an overview), the following ERP components were used for analysis in each task: the N200 component was analysed between 200 and 400ms for the Flanker task. For the colour-shape switch task, the N200 and the P300 component were used, analysed in the time windows of 200-320ms (N200) and 280-400ms (P300). For the 2-back and 3-back task, the P300 component was analysed in the time window 280-400ms.

ERP latency is measured by computing the peak latency. The peak latency is determined by finding "[…] the maximum amplitude within a time window and use the latency of this peak as a measure of the latency of the underlying ERP component" (Luck 2005:237-238). The difficulty of this measurement is that that the measurement time window might show the increase or decrease of another large component at the edges of the time window which will displace the maximum or minimum peak. By using local peak latency measures rather than absolute latency peak measures, this problem can be resolved. In local peak latency measurement, "[…] a point is not considered a peak unless the three to five points on each side of it have smaller values" (Luck 2005:238). It must be noted that peak latency measures can be sensitive to noise which makes it all the more important to filter out noise before analysing the latency. Furthermore, it is sometimes not easy to determine when one component ends and another begins. Therefore, peak latency was analysed only within the specified time-window of the ERP component. The minimum (in case of the N200 component) or maximum (in case of the P300 component) peaks were computed and their latency extracted with MATLAB.

Frequency analysis was performed with MATLAB as well.

8.2 Results

This section covers results from different sources. First of all, competence data that examines changes in different aspects of interpreting competence indexed by the TICQ introduced in Section 8.1.3. Subsequently, behavioural results including reaction times and accuracy of all three executive function tasks are presented. This is followed by the presentation of electrophysiological results, initially in the form of event-related potentials of different ERP components and the voltage changes and differences between conditions in relevant ERP component time-windows. Finally, data from one participant was chosen for the presentation of time-frequency results across all tasks and sessions.

8.2.1 Competence data

Competence data from the interpreting group is analysed by extracting data from the TICQ that participants filled out before taking part in the study and after taking part in the study at the end of their four-semester interpreting training. For general comparison between groups and comparison of pre- and post-questionnaire answers, the raw data from the respective TICQ questions was analysed.

Participants rated their English proficiency with a mean of 76.11 points out of 100 (SD = 12.44) in the pre-test and with a mean of 78.75 points out of 100 (SD = 10.26) in the post test. Although the mean L2 proficiency rating is higher in the post-test, a t-test did not show a significant difference between pre- and post-questionnaire administration (t = 0.42097, df = 13.157, p-value = 0.6806).

In terms of their self-rated proficiency to interpret from their L1 into their L2 (English), results showed a mean of 44.44 points out of 100 in the pre-test (SD = 19.6) and a mean of 58.38 points out of 100 in the post-test (SD =31.1). However, a t-test did not reveal a significant difference between groups t = (t = 0.85003, df = 11.411, p-value = 0.4128), probably due to one participant only indicating 1 from 100 points in the post-test, resulting in a high standard deviation. In the other direction, interpreting from L2 into their L1, results indicated a mean of 61.67 points out of 100 in the pre-test (SD = 18.20) and 75.25 points out of 100 in the post-test (SD =19.17). Although there clearly is an increase in mean from pre- to post-test, the t-test did not show a significant difference between pre- and post-session (t = 1.1255, df = 13.134, p-value = 0.2805).

Participants spent a mean of 4.25 hours per week interpreting from their L1 into their L2 (SD = 1.17) as indicated by the pre-test and 6.5 hours per week interpreting from their L1 into their L2 (SD = 3.82) as indicated by the post-test. The result of a t-test was not significant (t = 1.5946, df = 8.2927, p-value = 0.1481). In the other translation direction, from L2 into their L1, participants indicated that they spent on average 4.25 hours per week interpreting in this translation direction at the beginning of training (SD = 1.17) and on average 4.25 hours per week in this translation direction at the end of training (SD = 2.49). Results of a t-test confirmed the observation that participants did not significantly increase their weekly amount of interpreting from L2 into L1 over the course of the study (t = 0, df = 9.9183, p-value = 1).

Regarding the self-reported maximum simultaneous interpreting capacity per day, participants indicated a mean value of 1.67 hours (SD = 1.33) at the beginning of training and a mean value of 2.44 hours (SD = 1.29) at the end of their training. Although their capacity seems to have increased, results of the t-test did not reach significance (t = 1.1412, df = 13.987, p-value = 0.273).

Finally, participants indicated a mean décalage of 3.83 seconds at the beginning of their training (SD = 1.72) and a mean décalage of 2.44 seconds at the end of their training (SD = 1.29). Again, although the means have increased from pre- to post-test, t-tests results do not indicate a significant difference (t = -1.6639, df = 8.9791, p-value = 0.1306), probably because two participants did not indicate their décalage in the pre-questionnaire.

The subsequent part of the TICQ covers the preparation phase of simultaneous interpreting. Participants considered the translation brief before starting their preparation in 42.5% of the cases at the beginning of their training (SD = 46.83) and in 73.5% of the cases at the end of their training (SD = 43.04). Although this indicates a big leap from pre- to post-test, a t-test did not indicate a significant difference between testing sessions (t = 1.3786, df = 13.901, p-value = 0.1898).

Participants were asked whether they prepared printed glossaries, digital word lists or a digital term database for their interpreting task. Only 12.5% indicated that they prepared printed glossaries in the pre-questionnaire and 50% indicated that they did so in the post-questionnaire. None of the participants in the pre- or post-questionnaires indicated that they prepared a digital term database for their interpreting task, but 75% indicated that they used a digital word list at the beginning of training and 62.5% indicated that they still do so at the end of their training.

Of the participants, 62.5% indicated that they consulted external sources (websites, fora, databases, books) before the interpreting assignment at the beginning of their training and 87.5% of participants indicated that they do so at the end of their training. Similarly, 50% of participants indicated that they consult own resources (e.g., glossaries, translation memories, term databases) before the interpreting assignment at the beginning of their training and 87.5% indicated that they do so at the end of their training.

All participants at the beginning as well as at the end of their training indicated that they do not use translation memory systems or terminology databases for data management. Instead, they indicated that they use search engines such as Google where they predominantly looked for background information, idiomatic expressions, words, or phrases. None of the participants indicated that they use interpreter-specific tools, e.g., Interplex or Interpretbank. When asked in which language they google, results in the pre-questionnaire indicated that 37.5% google in the target language, 0% in the source language and 62.5% in both source and target language. In the post-questionnaire, this distribution did not change.

In terms or research aids, participants indicated using the following resources at the beginning of their training: bilingual dictionaries (87.5%), monolingual dictionaries (62.5%), synonym dictionaries (77.78%), encyclopaedias (37.5%), search engines (87.5%), terminology databases (12.5%) and concordance searches (12.5%). At the end of their training, usage in all resources increased (bilingual dictionaries: 100%; encyclopaedias: 75%; search engines: 100%, terminology bases: 12.5%; concordance search 25%) with the exception of monolingual dictionaries, which were used by 62.5% and synonym dictionaries which were used by 50%.

For the strategies applied during interpreting, participants rated the importance of visual contact to the speaker during simultaneous interpreting. In the beginning of their training, participants indicated a mean of 47.5 points out of 100 (SD = 32.84) and at the end of their training they indicated a mean of 57.5 out of 100 points (SD = 36.15). A t-test did not reveal a significant difference between sessions (t = 0.57908, df = 13.873, p-value = 0.5718). Furthermore, participants were asked what strategy they used when they do not understand the speaker acoustically. In the pre-questionnaire, 62.5% indicated omitting the passage they did not understand acoustically, 12.5% indicated inventing new content and 25% did not indicate a strategy at all. In contrast to that, 50% of participants in the post-questionnaire indicated that they omit the passage they did not understand acoustically, 12.5% asked their partner, 12.5% made a guess and 0% indicated that they invent new content. Still, 25% did not indicate a strategy.

Participants were also asked for their strategy when they do not understand the meaning of a term or sentence during the interpreting process. In the pre-questionnaire, asking a partner was not represented at all while 62.5% indicated doing so in the post-questionnaire, omission was initially represented with 37.5% and increased to 62.55% in the post-test, guessing the content initially lay at 62.5% and at 50% in the post-questionnaire, inventing new content was initially represented with 25% and not represented at all in the post-questionnaire, interrupting the speaker was not an option for any participant in the pre- or in the post-questionnaire, looking something up in resources was only represented with 12.5% at the beginning of training but applied as an option by 50% at the end of training and finally, stalling was initially chosen by 50% and increased to 87.5% at the end of training.

Upon asking how the participants deal with spontaneous changes in the manuscript, if they get one, only 37.5% of participants initially indicated that they interpret the changes while 87.5% in the post-test opted for interpreting the changes. The remaining participants did not indicate a strategy.

In total, 50% of participants at the beginning of their training indicated that they correct obvious mistakes of the speaker in the target language which increased to 75% of participants at the end of training correcting obvious mistakes of the speaker in the target language.

Participants were also asked which resources they used during the SI process in the booth, which was similar to the question of which resources they prepared. Interestingly, none of the participants initially used online resources which increased to 75% use of online resources in the post-questionnaire. No term bases or translation memories were used at the beginning or at the end of training by any participant. Only 12.5% indicated using Google in the booth at the beginning of training which increased to 75% of use at the end of interpreting training. Printed glossaries were not so commonly used as Google, with only 37.5% indicating that they use printed glossaries at the beginning of training and 37.5% indicating use of printed glossaries at the end of training. Digital lists were used much more in the post-questionnaire, with 87.5% of participants indicating that they use them while only 12.5% used them at the beginning of training. Finally, interpreter-specific tools such as Interplex or Interpretbank were not used at all.

Conclusively, participants were asked about how they assessed the quality of their interpretation. Initially, 25% of participants indicated that they receive feedback from their colleagues, 75% indicated that they assess the interpreting quality through recordings. This shifted at the end of their training, where they did not rely as much on quality assessment through recordings (37.5%) and rather on feedback from their colleagues (37.5%) and clients (12.5%). Still, one participant indicated that they did not assess the quality of their interpretation at all.

Summed competence values were calculated per participant and values from the pre-training questionnaire were compared to values in the post-training questionnaire by means of a t-test. Table 7 shows the summed competence values for the SI competence section of the TICQ per participant from the pre-training and the post-training questionnaire.

Table 7: Summed TICQ values per participant for pre- and post-training questionnaire

	Pre	Post
P01	3.018	4.83
P03	4.2365	5.746
P04	3.061	4.698
P06	6.039	5.579
P07	4.8605	5.95
P08	4.53	6.135
P10	4.0195	4.734
P11	4.612	3.25525

The mean pre-training competence value was 4.297 (SD = 0.919) and mean post-training competence value was 5.116 (SD =0.881). A t-test confirmed the hypothesis that overall competence increases from the beginning of training to the end of SI training, since it indicated a significant difference between pre- and post-training competence value (t = 1.8945786, df = 7, p = 0.0415). To identify the strength of the effect, the Cohen's d was calculated by subtracting mean 2 from mean 1 and dividing it through the pooled standard deviation of both means. Results revealed a Cohens' d value of 0.9098 which can be classified as a strong positive effect, indicating that competence increased significantly from the pre-training questionnaire to the post-training questionnaire. This result is concurrent with other studies that have found gains in interpreting performance after interpreting training (see Tzou et al. (2012), Hervais-Adelman et al. (2015a) and Yu & Dong (2022)).

The distribution of the F1-values for each participant for the pre- and post-questionnaire are depicted in Figure 6. While most participants seem to have increased their competence, there are two participants, P06 and P11, who do not show a higher F1-value in the post-questionnaire than in the pre-questionnaire. Also striking is that P06 started out with the second highest value of all participants in both questionnaire sessions. These outliers might represent the caveat of self-rated questionnaires or a general difficulty to rate one's competence after having learned more about the skill that is acquired, in this case simultaneous interpreting.

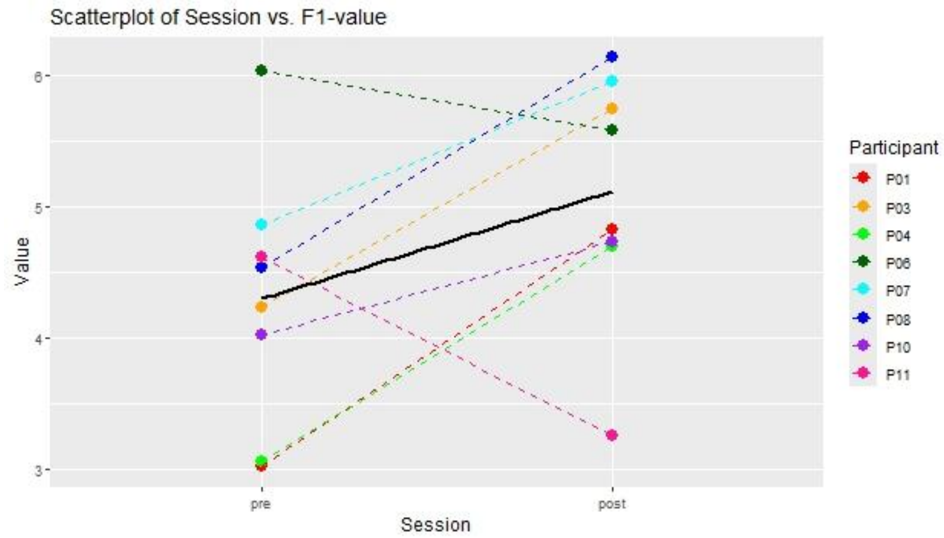


Figure 6: Scatterplot of F1-values per participant in pre- and post-questionnaire

To not only compare the means of two groups, I furthermore fitted a linear mixed model with the F1-value as dependent variable, Session as predictor and Participant as random effect. The effect of Session on the F1 value revealed only a certain trend toward significance (p = 0.0832, SE = 0.4054); the effect was medium-sized as indicated by Cohen's f^2 ($f^2 = 0.19$) with the fixed effects coefficient of the post-questionnaire increasing by 0.8188 units (see Figure 7). The statistically insignificant p-value could possibly be due to the two participants that did not increase their self-rated interpreting competence. These SI competence results will be descriptively triangulated with the ERP results as well as the behavioural results in the Discussion in Chapter 9.

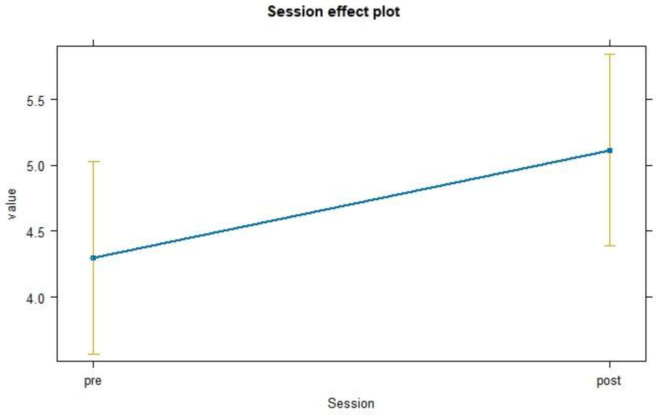


Figure 7: Linear mixed model of F1-values per Session

Lastly, I would like to mention the attempt to also collect objective competence data for this study. By obtaining written consent from all participants, I was able to get exam results from two different simultaneous interpreting exams. The first exam took place at the end of the first semester of testing and constituted a "mock exam" where the students needed to simultaneously interpret a text from English into their L1. They were evaluated on the basis of the standard simultaneous interpreting evaluation sheet of the Master of Conference Interpreting that is also used in their final exam. An example is attached in the Appendix. The "mock exams" were not graded and are meant to help the participants to evaluate their own performance. The last exam constitutes the final exam of the course of studies and took place at the end of the fourth semester, the last semester of testing.

As can be seen from an example sheet in the Appendix, the interpreting performance was divided into 4 sections. Section 1 covered the presentation of the interpreting performance, including points for pronunciation, intonation, completion of sentences, communicative competence, and fluency. Section 2 covered the content of the interpreting performance, containing completeness, sense, cultural competence, coherence, and monitoring. Section 3 dealt with the language during interpreting, covering grammar, syntax, and expression. Lastly, Section 4 covered whether or not the students successfully applied interpreting strategies, however, there was no specification as to which strategies needed to be applied. The examiner put crosses down for the respecting category in the columns "good", "average" and "bad" performance. I initially planned to annotate the exam sheets and giving 3 points for the category "good", 2 points for the category "average" and 1 point for the category "bad" performance. Additionally, I decided to give 1.5 points when the examiner placed the cross between "bad" and "average" and 2.5 points when the examiner placed the cross between "average" and "good". Unfortunately, there were several problems with the evaluation sheets. First of all, an error occurred during the transfer process from the lecturer that conducted the evaluation to the university's examination office and as a result, I only received 8 out of 10 evaluation sheets. Two evaluation sheets could not be retrieved. Furthermore, not all students that participated in my study decided to take the final exam at the end of the fourth semester. Of 8 participants that remained in the study until the end of testing, only 4 took the final exam. Of one of these 4 participants, the initial exam at the end of the first semester was lost and no pre- and post-training evaluation could take place. This leaves me with objective competence data from three participants (P03, P04 and P07) that can be evaluated and is described in the following.

As explained, I annotated the exam evaluation sheet according to the above mentioned grading system. Participants could obtain a maximum of 42 points. Table 8 shows the summed points from the annotation of the initial and the last exam for P03, P04 and P07.

Table 8: Overview of points from initial as well as last SI exam

	Initial Exam	Last Exam
P03	40	19.5
P04	37	39
P07	27	25

What is striking is that P04 is the only participant that showed a slight increase in interpreting competence. Her biggest gains lay in the improvements of translating the sense of what was uttered by the speaker as well as gains in cultural competence, expression, and application of interpreting strategies. P07 even performed slightly worse in her final exam, loosing points in applying interpreting strategies, syntax, grammar as well as all aspects of presentation. P03's performance in the last exam was quite poor compared to her initial exam which led to P03 not passing the final exam at all. Although these results from only three participants are merely anecdotal and cannot be a part of the overall analysis, the have some interesting implications about interpreting competence that will be further expanded upon in the Discussion (Chapter 9).

8.2.2 Behavioural data

For the behavioural data, reaction times as well as accuracy rates were analysed between the Experimental as well as the Control group at baseline level as well as for the longitudinal data of the Experimental group for all tasks. Accuracy rates were calculated by dividing the number of correct responses by the number of total responses. For the between-group comparison of reaction times at baseline level in all tasks, linear mixed models were calculated with reaction time as outcome variable, Group and Condition as predictor variables as well as Participant as random effect.

8.2.2.1 Flanker task

Baseline between-group comparison

Reaction times for the Experimental and the Control group were compared at baseline level to see if there is an initial advantage in reaction times for the Interpreting group (Experimental group). A linear mixed model fitted with reaction time as outcome variable, Group and Congruency as predictor variables and Participant as random effect revealed no significant effect of Group on reaction time ($p = 0.693$, SE =1.765e-02). The effect of Congruency on reaction time was highly significant ($p < 0.01$, SE =1.975e-03) and showed a large effect size ($f^2 = 0.33$) with the incongruent condition having significantly slower reaction times (see Figure 8). VIF-values were equal to 1 and therefore did not indicate collinearity.

Visual inspection of a Q-Q plot for the residuals showed that residuals deviated from normality. Therefore, the linear model was log-transformed to make the residuals more normally distributed. The log transformed linear model did not show remarkable differences to the original linear model.

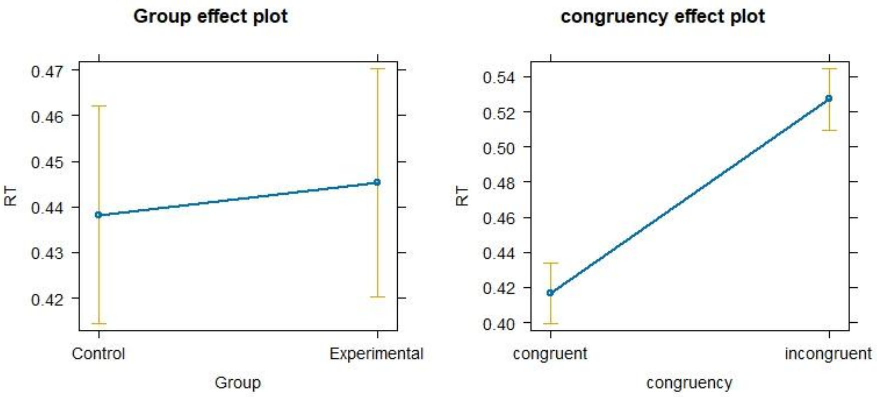


Figure 8: Linear mixed model of between-group comparison of Flanker RT at Baseline

After including an interaction term to see whether Condition is modulated by Group, results revealed a very small ($f^2 < 0.01$) but significant interaction of Condition and Group (p < 0.001, SE = 3.945e-03) that was driven by the incongruent condition, showing that the Experimental group displayed slower reaction times especially in the incongruent condition (see Figure 9).

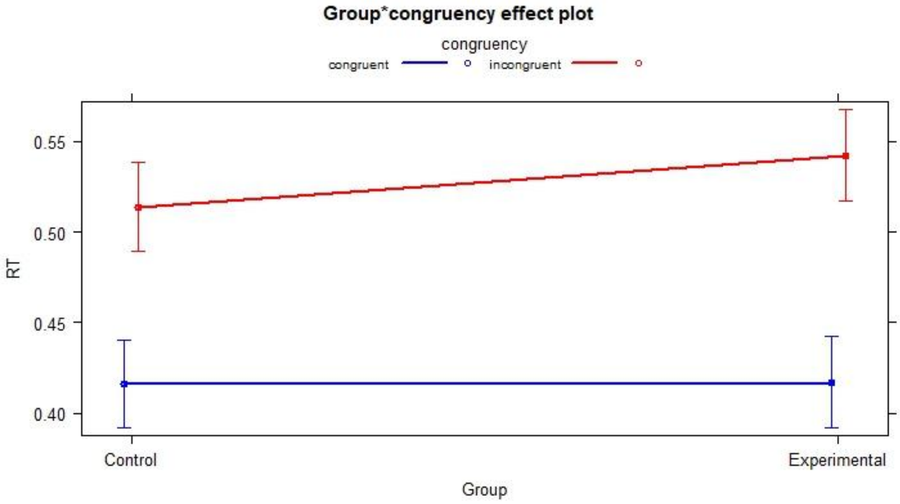


Figure 9: Interactions Group and Condition for reaction times at Baseline

Mean reaction times of congruent and incongruent conditions for both groups as well as the differences between conditions for both groups and accuracy per condition are depicted in Table 9. A separate linear model for the effect of Group on reaction time in the congruent condition did not reveal a significant difference between groups (p =0.963, SE =1.843e-02), nor did a separate model for the incongruent condition (p =0.187, SE =0.02142).

Table 9: Descriptive statistics for Baseline between groups comparison

	Experimental group	Control group
Mean congruent RT (in ms)	0.417, SD = 0.078	0.416, SD =0.074
Mean incongruent RT (in ms)	0.541, SD = 0.098	0.513, SD = 0.084
Difference (in ms)	0.124	0.097
Overall accuracy	0.9530	0.956
Accuracy congruent	0.9881, SD = 0.108	0.983, SD = 0.130
Accuracy incongruent	0.8477, SD = 0.359	0.876, SD = 0.330

Accuracy was compared between both groups by fitting a logistic regression model with Accuracy as the outcome variable, Group and Congruency as predictor variables and Participant as random effect to account for differences in accuracy between groups (see Figure 10). Results did not a show a significant effect of Group on accuracy (p = 0.517, SE = 0.11436) although the Experimental group decreased by -0.06930 units. However, the effect of Congruency was large (f^2 = 0.32) and highly significant (p < 0.01, SE = 0.12918), with incongruent condition decreasing by -2.37387 units, showing that incongruent trials yielded significantly more error than congruent trials. The VIF-value was equal to 1 and therefore did not indicate collinearity.

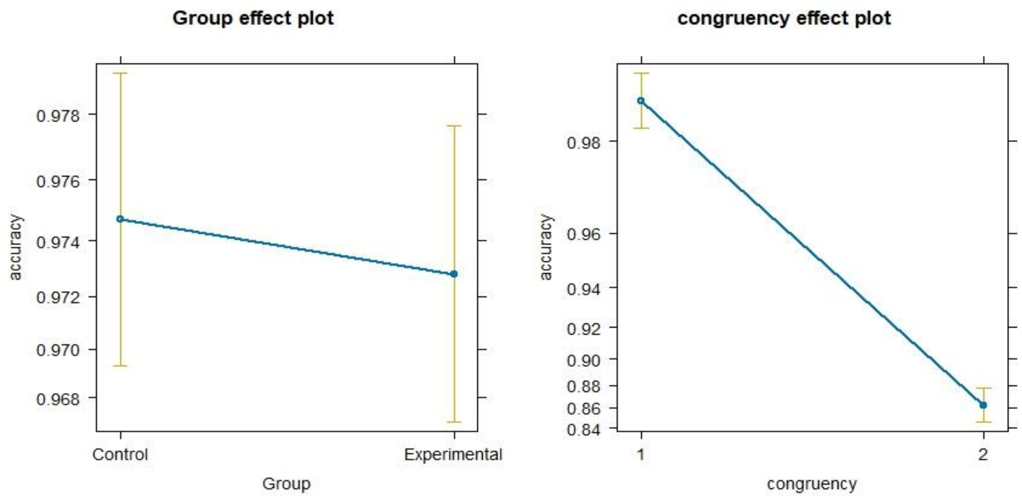


Figure 10: Linear mixed model of accuracy in Control and Experimental group at Baseline

The inclusion of an interaction term to see whether Congruency is modulated by group revealed a significant interaction between Group and Congruency (p = 0.0203, SE = 0.2636), showing that the Experimental group had lower accuracy in the incongruent condition (see Figure 11). The Group:Congruency interaction effect size was small ($f^2 = 0.01$).

The main effects of group in both RT and accuracy that do not show differences between groups at baseline correspond to previous studies that did not find an SI advantage in inhibition at baseline testing such as by Babcock et al. (2017), Dong & Liu (2016) and Rosiers et al. (2019). The main effects of Congruency in RT and accuracy, showing faster RT and higher accuracy for congruent than for incongruent trials also correspond to previous literature (see Verreyt et al. (2016), Luk et al. (2010), Dong & Xie (2014), van der Linden et al. (2018)).

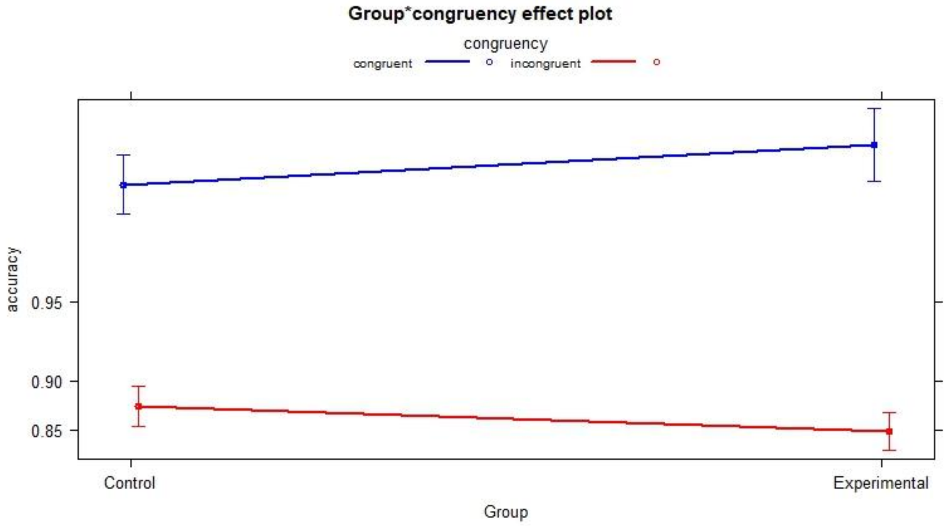


Figure 11: Interaction between Group and Condition for accuracy at Baseline

Longitudinal results of Experimental group

Figure 12 depicts the progression of the mean reaction time for the Experimental group across all sessions. Although reaction times of both conditions decrease steadily, reaction times of the incongruent condition always remain longer than reaction times in the congruent condition.

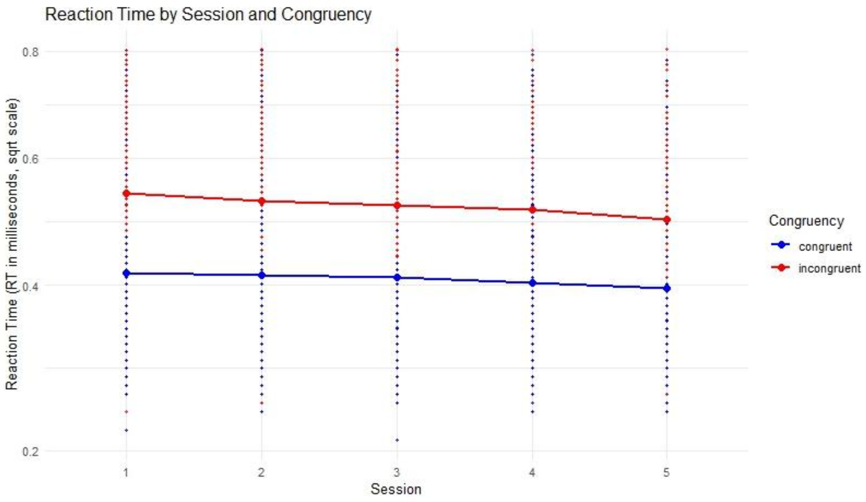


Figure 12: Progression of mean reaction time per session and condition

This trend is confirmed by a linear mixed model fitted with reaction time as outcome variable, Session and Congruency as predictor variables and Participant as random effect. The effect of all Sessions on reaction time was significant (S2: $p = 0.0129$, SE = 1.790e-03; S3: $p < 0.01$, SE = 1.793e-03; S4: $p < 0.01$, SE = 2.059e-03; S5: $p < 0.01$, SE = 1.929e-03). The overall effect size of Session was small ($f^2 = 0.02$). Furthermore, the model showed a large ($f^2 = 0.36$) and highly significant effect of Congruency on reaction times ($p < 0.01$, SE = 1.426e-03) with incongruent trials showing significantly longer reaction times (see Figure 13). The VIF-values for both Session and Congruency were equal to 1, not indicating any multicollinearity.

Visual inspection of the residuals of the linear model fitted above showed deviation from normal distribution. Therefore, the linear model was log transformed to check whether this changes the effects. Only marginal changes were found after log transformation that did not alter the overall results.

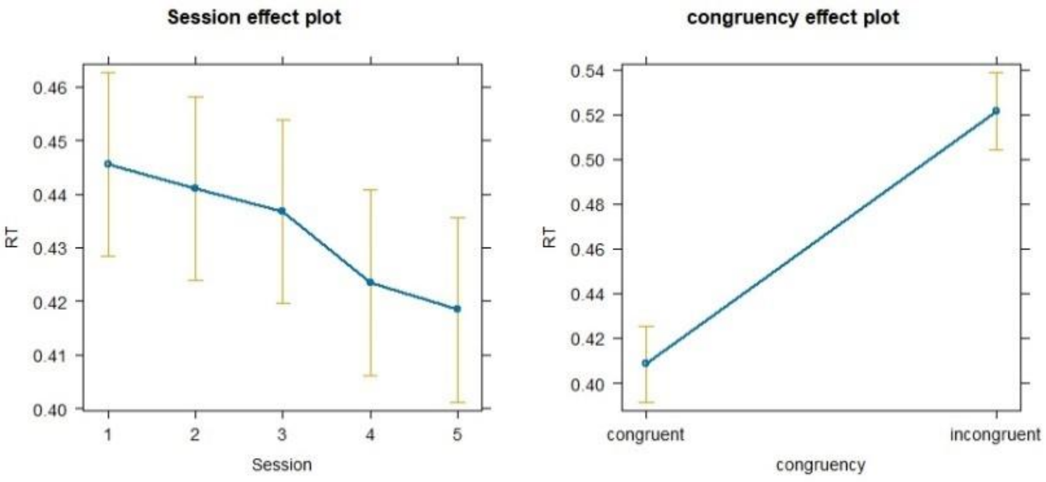


Figure 13: Linear mixed model of reaction time progression across all sessions

Moreover, an interaction term of congruency was added to the linear mixed model to test whether Condition influences how Session affects the reaction time, as depicted in Figure 14. Results showed significant interactions between all Sessions and Congruency, more specifically with incongruent trials, indicating that the effects on Session 2-5 on reaction time is different in the incongruent condition compared to the congruent condition (S2: p= 0.010955, SE = 4.280e-03; S3: p = 0.000529, SE = 4.290e-03; S4: p = 0.002370, SE = 4.770e-03, S5: p < 0.01, SE = 4.512e-03). Results indicated that especially the RT in the incongruent condition decreased as Sessions progressed. The overall effect size of the Session:Condition interaction was small (f^2 < 0.01).

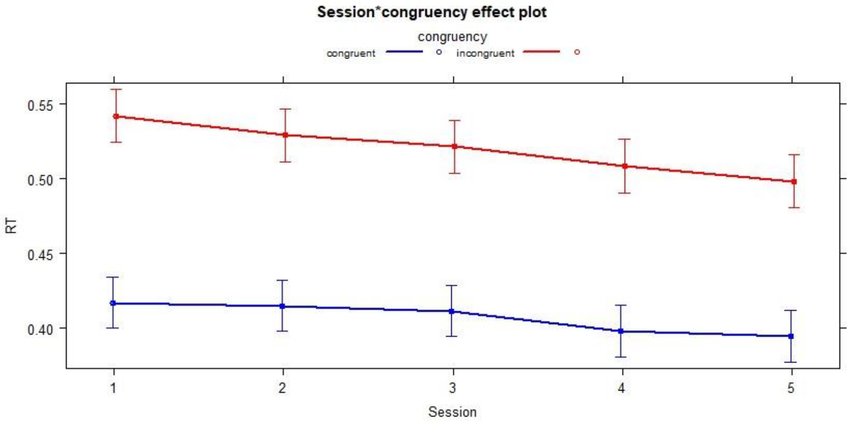


Figure 14: Linear mixed model with congruency as interaction term

I additionally applied the emmeans() function in R that estimates marginal means for specified factors in linear models using emms <- emmeans(mixed_lm2, ~ Session * congruency) and the pairwise comparison of Conditions using emmeans(mixed_lm2, pairwise ~ congruency | Session). This way, I checked for the pairwise difference between the conditions across Session progression. Results showed that that the difference between conditions decreased as Session progressed, indicated by the coefficients in Table 11 and Table 10:

Table 10: Difference between means per session

Table 11: Emmeans per Session and Condition

Session1	**emmeans**	**SE**
congruent	0.417	0.00876
incongruent	0.542	0.00905
Session2		
congruent	0.415	0.00876
incongruent	0.529	0.00904
Session3		
congruent	0.412	0.00876
incongruent	0.521	0.00904
Session4		
congruent	0.398	0.00884
incongruent	0.509	0.00923
Session 5		
congruent	0.395	0.00880
incongruent	0.498	0.00913

Session1	**estimate**	**SE**	**p-value**
contrast	-0.125	0.0030	<.0001
Session2			
contrast	-0.114	0.00301	<.0001
Session3			
contrast	-0.110	0.00302	<.0001
Session4			
contrast	-0.110	0.00367	<.0001
Session 5			
contrast	-0.104	0.00333	<.0001

The values for the Flanker effect were obtained by calculating the difference of the mean reaction times between conditions per session. As depicted in Figure 15, the Flanker effect or more precisely the influence of the incongruent stimuli decreased until Session 3, after which it slightly increased for Session 4, followed by another decrease after Session 5.

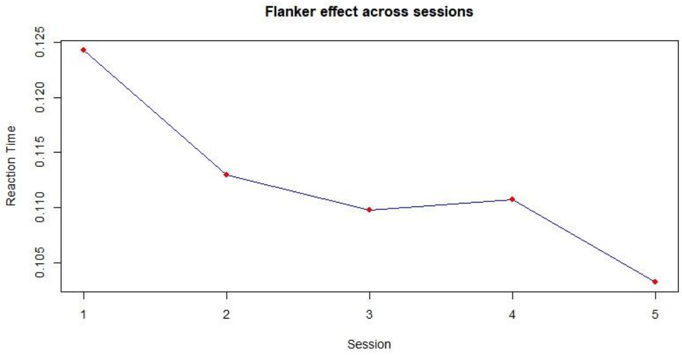


Figure 15: Flanker effect across sessions

As depicted in Figure 16 (left), accuracy initially increased but then reached a negative peak at Session 4, after which it increased again drastically. Taking a closer look, Figure 16 (right) shows that the congruent condition increased more drastically after baseline, while the negative peak in Session 4 as well as the drastic increase is mirrored in both conditions.

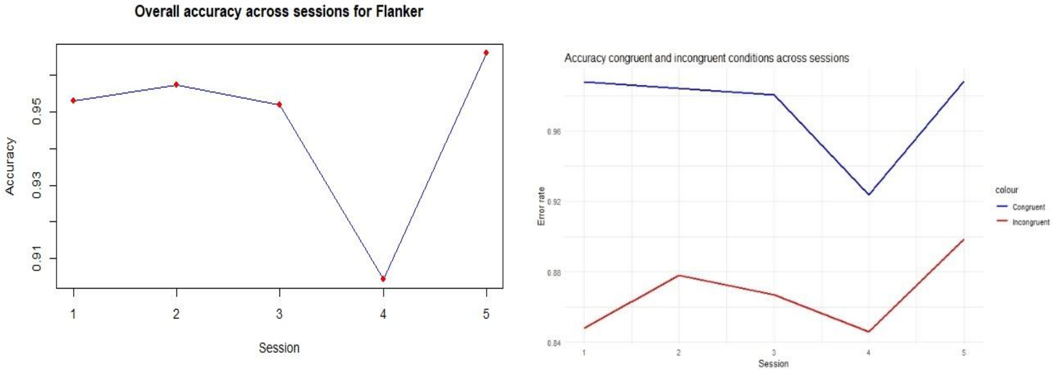


Figure 16: Accuracy across sessions for the Flanker task

A logistic regression model with Accuracy as outcome variable and Session and Condition as predictors revealed that Session 2 and Session 3 did not show significant effects on Accuracy (S2: $p = 0.32584$, SE = 0.11673; S3: $p = 0.82080$, SE = 0.11327), while Session 4 and Session 5 showed highly significant effect of Session on Accuracy (S4: $p < 0.001$, SE = 0.10716; S5: $p = 0.00832$, SE = 0.13186). The overall effect size of Session was small ($f^2 < 0.01$). The effect of Congruency on Accuracy was medium-sized (f^2 = 0.19) and highly significant (p < 0.001, SE = 0.07540) with the incongruent condition decreasing by -1.83165 units from the congruent condition, indicating significantly lower accuracy for the incongruent condition. VIF-values for both Session and Condition were equal to 1, not pointing towards collinearity.

Adding an interaction term to see how Session modulates Condition (see Figure 17) revealed a marginally significant interaction for Session 2 (p = 0.0530, SE = 0.2735) and significant effects for Sessions 3 and 4 (S3: p = 0.0112, SE = 0.2629; S4: p < 0.001, SE = 0.2448). The effect size for the Session:Condition interaction was small (f^2 = 0.09).

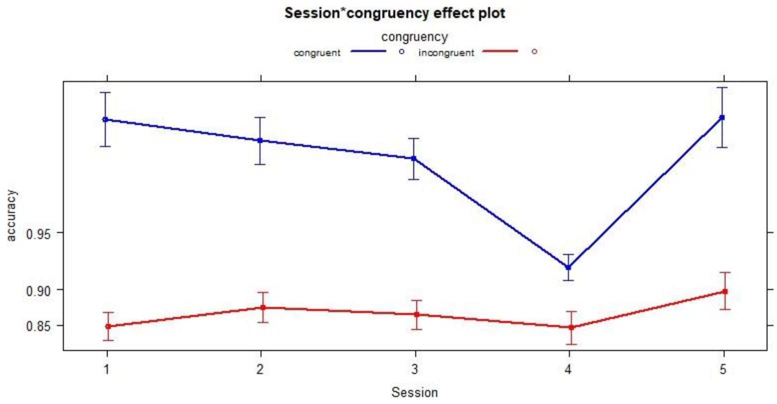


Figure 17: Interactions Condition and Session for accuracy

The effect seemed to be driven by the incongruent condition, with accuracy increasing more in the incongruent condition than in the congruent condition over the course of testing as depicted by the estimated marginal means and the contrasts between conditions in Table 12 and Table 13. The massive decrease in accuracy in represented by the negative coefficient of Session 4 indicated in the interaction results in R (estimate: -1.9281, SE = 0.1978, p < 0.001).

Table 12: Estimated marginal means for accuracy across session and condition

Table 13: Contrast of emmeans for accuracy across session and condition

Session1	emmeans	SD
congruent	4.42	0.1778
incongruent	1.72	0.0928
Session2		
congruent	4.15	0.1555
incongruent	1.97	0.1018
Session3		
congruent	3.91	0.1387
incongruent	1.87	0.0981
Session4		
congruent	2.50	0.0867
incongruent	1.70	0.1104
Session5		
congruent	4.45	0.2012
incongruent	2.18	0.1235

Session1	estimate	SE	p-value
contrast	2706	0.201	<.0001
Session2			
contrast	2177	0.186	<.0001
Session3			
contrast	2039	0.170	<.0001
Session4			
contrast	0.791	0.140	<.0001
Session5			
contrast	2265	0.236	<.0001

The longitudinal results that RT in the Flanker task sped up over testing correspond to previous results of Babcock et al. (2017) who observed faster reaction times in the monitoring section of the Attention Network Task after 2 years of SI training. However, this effect was not SI specific, since all tested participant groups showed faster RT at the post-test compared to the pre-test. Similarly, Dong & Liu (2016) who used a Stroop task found faster RT across all groups after one semester of respective interpreting, translation or language training. This effect was therefore also not SI specific.

The present results show that participants did increase their overall accuracy in reacting to congruent and incongruent trials indicated by the significant effect of S5 on accuracy, an effect not observed in the longitudinal studies by Dong & Liu (2016) and Babcock et al. (2017).

8.2.2.2 Colour-shape switch task

Baseline between-group comparison

Between-group comparison of reaction times at baseline level for the colour-shape switch task (see Figure 18) did not reveal a significant effect of Group (p = 0.984, SE = 1.078e-01), indicating no initial advantage for the interpreting group in terms of overall reaction times. The effect of Condition was small ($f^2 = 0.02$) but significant (p < 0.01, SE = 1.044e-02) with an increase of 1.158e-01 units for the switch condition, showing that the overall reaction time for switch trials were significantly larger than for repeat trials. Visual inspection of the residuals revealed that the data deviated from normality. Therefore, a log transformed linear model was fitted, which did not reveal substantially different results compared to the non-log transformed model. A VIF was calculated for this model; the VIFs for Group and Condition were equal to 1, indicating a low risk of collinearity.

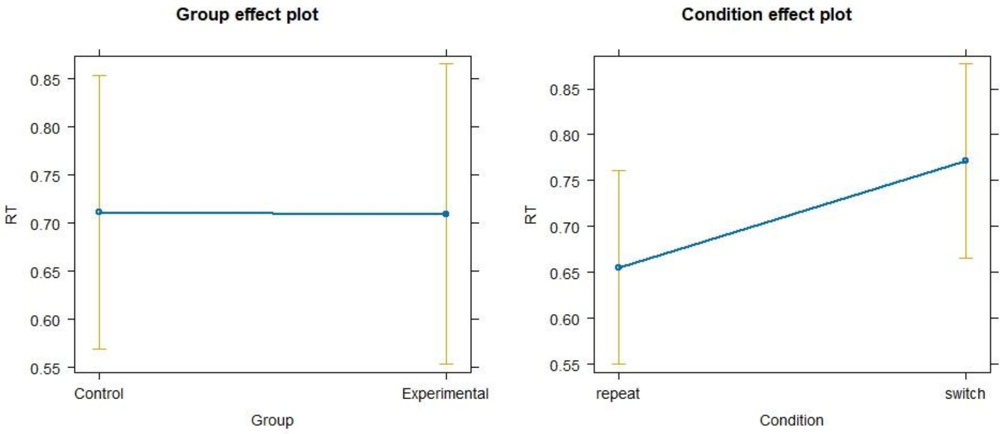


Figure 18: Linear mixed model of reaction times between groups at baseline

In terms of interactions between Condition and Group, results did not show a significant interaction (p = 0.075, SE = 0.02106) with only a tendency for shorter reaction times in the switch condition for the Experimental group compared to the Control group (see Figure 19)

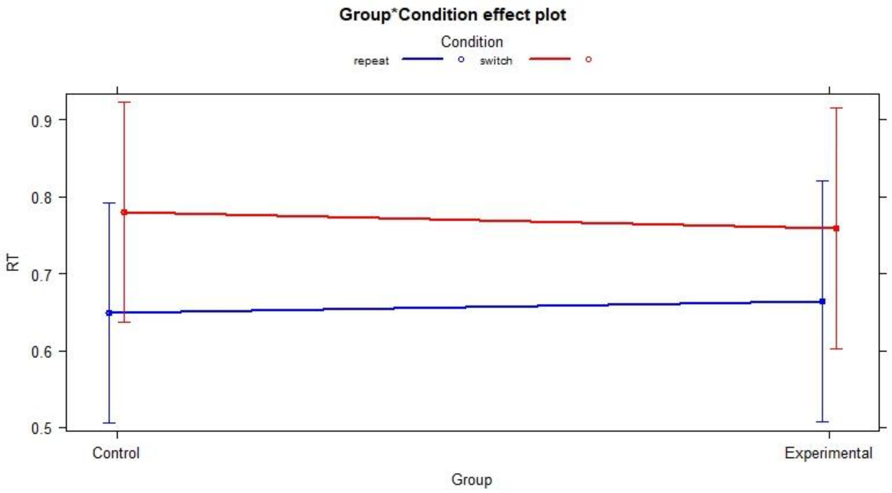


Figure 19: Interactions Group and Condition at baseline

Taking a closer look at reaction times at baseline between groups, results show the following mean reaction times for switch and repeat conditions per group as well as switching costs (calculated as the difference between repeat and switch trials) and accuracy in Table 14. The Experimental group showed smaller switching costs at baseline than the Control group.

Table 14: Descriptive statistics of both groups across Sessions in CSM

	Experimental group	Control group
Mean switch RT (in ms)	0.745, SD = 0.406	0.773, SD = 0.478
Mean repeat RT (in ms)	0.654, SD = 0.367	0.646, SD = 0.407
Switching costs (in ms)	0.091	0.128
Overall accuracy	0.868	0.928
Accuracy switch	0.842, SD = 0.364	0.914, SD = 0.280
Accuracy repeat	0.893, SD = 0.310	0.963, SD = 0.189

Separate linear mixed models for repeat and switch conditions between groups revealed that there is no significant difference between groups in repeat trials (p = 0.851, SE = 0.09577) or switch trials (p = 0.838, SE = 0.12502).

Overall, accuracy was high for both groups, however, it differed significantly between the groups as indicated by a logistic regression model with Accuracy as response variable and Group as predictor variable ($p < 0.01$, SE = 0.1033). This effect was small as indicated by Cohen's f^2 ($f^2 = 0.05$). Accuracy in the Experimental group decreased by -0.8571, indicating that the Experimental group made significantly more errors at baseline than the Control group. Furthermore, the effect of Condition on accuracy was small ($f^2 = 0.03$) but highly significant as well ($p<0.01$, SE =0.1024) with the switch condition decreasing by -0.6015 units, showing that significantly more errors were made in the switch condition (see Figure 20). The VIF for Group and Condition was equal to 1, indicating a low risk of multicollinearity.

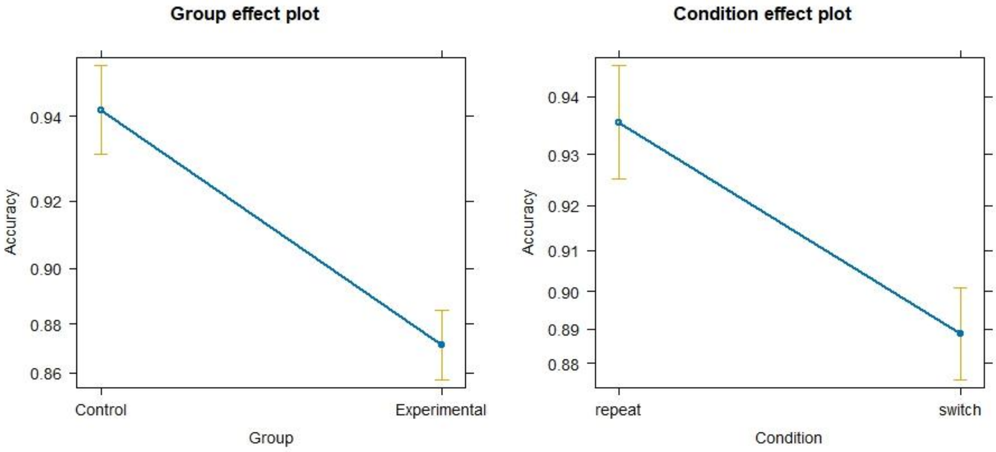


Figure 20: Logistic regression model for Accuracy at baseline between groups

Checking for interactions between Group and Condition by adding an interaction term revealed a significant interaction ($p = 0.0394$, SE = 0.2166) that seems to be driven by the switch condition (see Figure 21). The effect size of the interaction was small ($f^2 = 0.02$).

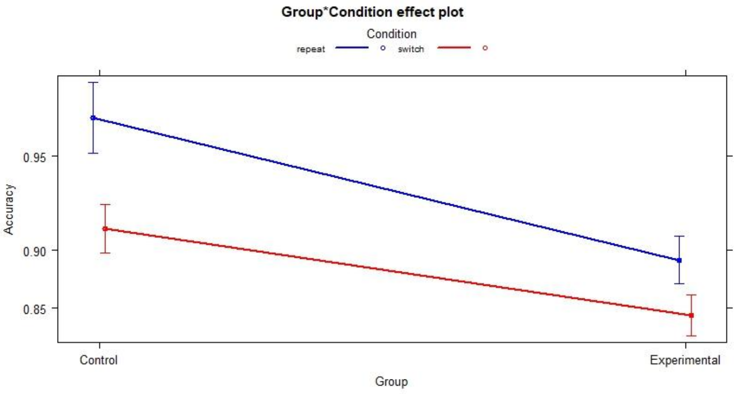


Figure 21: Interaction Group and Condition at baseline

The result that there are no differences between groups in reaction times at baseline in the colour-shape switch task correspond to results by Babcock et al. (2017), Dong & Liu (2016) and Rosiers et al. (2019) who all applied the colour-shape switch task and did not find a difference in RT between groups at baseline testing. The finding of a main effect of Congruency on RT with switch trials having lower RT than repeat trials corresponds to relevant task switching literature (see e.g. Garbin et al. (2010), Baene et al. (2015), Prior & Gollan (2011), Declerck et al. (2017).

The result that participants do differ in accuracy at baseline but that this difference is due to the interpreting group making more errors than the control group does contradict the findings by Rosiers et al. (2019) and Babcock et al. (2017) who did not find a difference in accuracy in the colour-shape switch task between groups, however, since the effect was driven by lower accuracy in the experimental group, no inherent advantage of the SI group at baseline is given.

Longitudinal results of Experimental group

Concerning the progression of the mean reaction time of the interpreting group that was tested longitudinally, Figure 22 shows that mean reaction time decreases steadily over the semesters, with mean switch reaction times always being slightly longer than mean repeat reaction times.

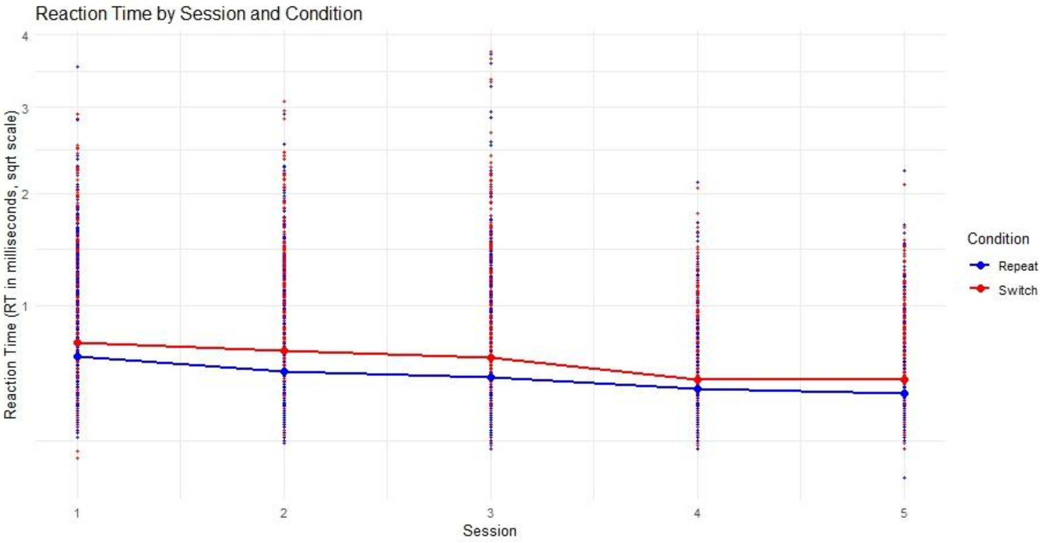


Figure 22: CSM progression of mean reaction time for switch and repeat condition

A linear mixed model (see Figure 23) with reaction time as outcome variable, Session and Condition as predictor variables as well as Participant as random effect shows significant effects for all sessions (Session2: $p < 0.01$, SE = 9.771e-03, Session 3: $p < 0.01$, SE = 9.461e-03, Session 4: $p < 0.01$, SE = 1.054e-02, Session 5: $p < 0.01$, SE = 1.018e-02). The general effect size of Session was small as indicated by Cohen's f^2 (f^2= 0.07). Furthermore, the effect of Condition on reaction time was also small (f^2 = 0.07) but highly significant ($p < 0.01$, SE = 6.270e-03) with the switch condition showing slower reaction times than the repeat condition. Plotting the residuals revealed that residuals are 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. The VIF-values were equal to 1 for both Session and Condition which indicated that there is no collinearity between independent variables.

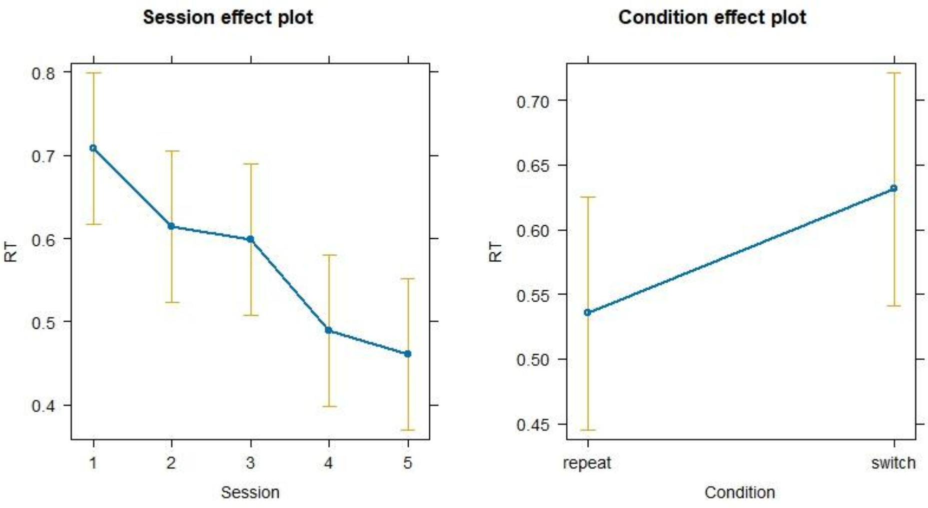


Figure 23: Linear mixed model of reaction time across sessions

Finally, Condition was added as an interaction term to the linear mixed model, to find out whether the effect of Session on reaction time is modulated by Condition (see Figure 24).

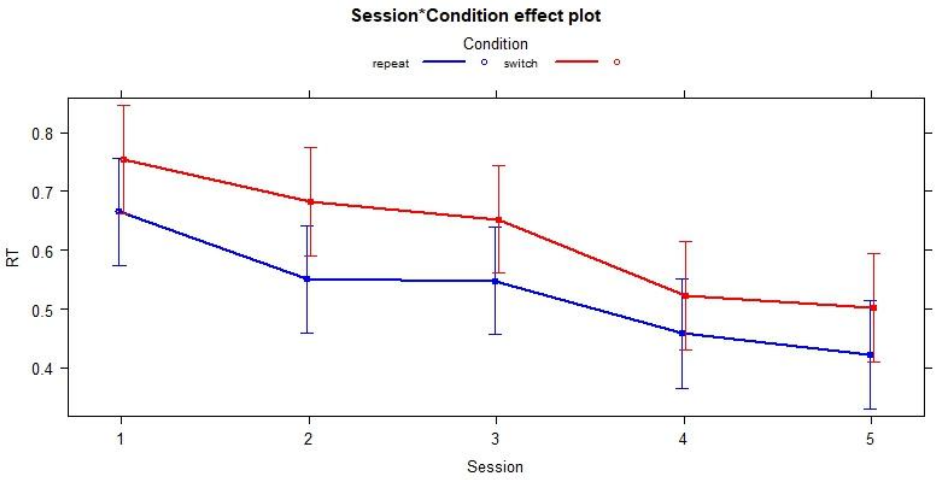


Figure 24: Interaction term in linear mixed model

The linear mixed model only revealed a significant interaction with Session 2 (p =0.0254, SE =1.928e-02) that was driven by the switch condition, showing a greater difference in reaction times between congruent and incongruent condition in Session 2 than Session 1 as indicated by pairwise comparison of estimated marginal means in Table 16 and Table 15. The effect size of the Session:Condition interaction was extremely small ($f^2 < 0.01$).

Table 15: Contrasts between means per session

Table 16: Emmeans of each condition per session

Session1	**emmeans**	**SE**
repeat	0.665	0.0467
switch	0.755	0.0468
Session2		
repeat	0.551	0.0467
switch	0.683	0.0468
Session3		
repeat	0.548	0.0466
switch	0.652	0.0467
Session4		
repeat	0.458	0.0471
switch	0.523	0.0471

Session1	**estimate**	**SE**	**p-value**
contrast	-0.0893	0.0136	<.0001
Session2			
contrast	-0.1324	0.0137	<.0001
Session3			
contrast	-0.1043	0.0132	<.0001
Session4			
contrast	-0.0650	0.0154	<.0001
Session5			
contrast	-0.0800	0.0145	<.0001

Session 5		
repeat	0.422	0.0469
switch	0.502	0.0470

Behavioural switching costs were calculated per session for the longitudinally tested interpreting group based on the difference in reaction times between switch and repeat trials. As depicted in Figure 25, switching costs initially increased, then decreased and hit a negative peak at Session 4.

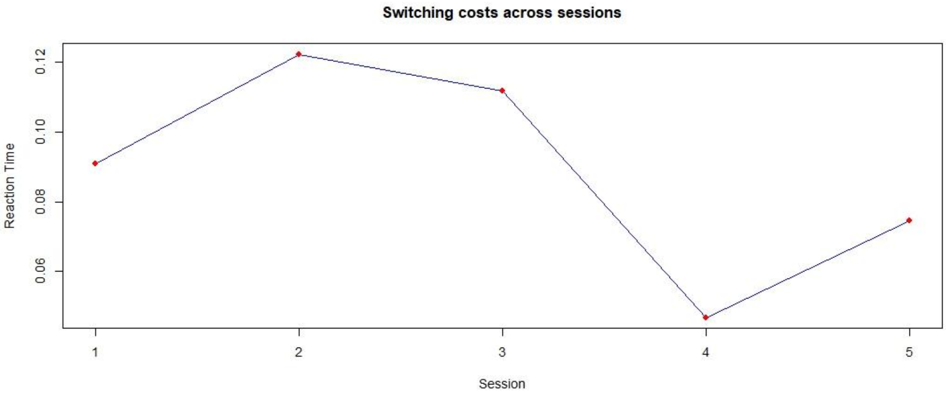


Figure 25: Progression of behavioural switching costs across sessions

As depicted in Figure 26, 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 $p > 0.05$).

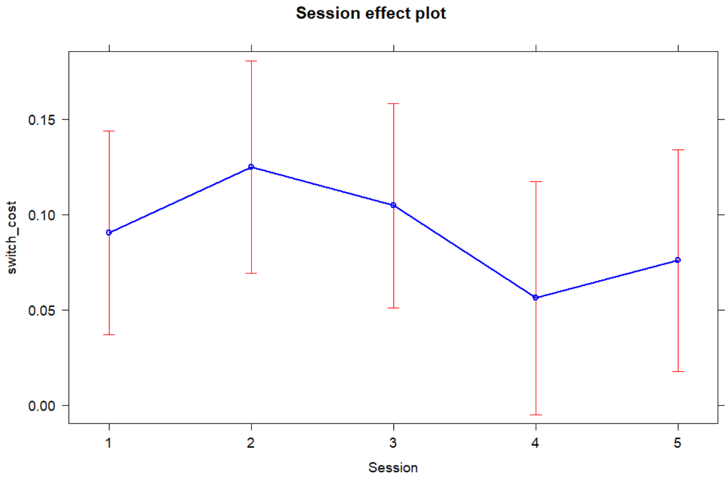


Figure 26: Linear mixed model of switch costs across all sessions

As depicted in Figure 27, the overall accuracy of the interpreting group across the sessions increases drastically from Baseline to Session 2 but experiences a decrease at Session 3 after which it increases again for Session 4, followed by another slight decrease in Session 5.

Figure 28 depicts the progression of accuracy for both conditions separately, showing a clear synchronicity between the conditions except for Session 5, where accuracy in repeat conditions increases while accuracy in the switch condition decreases.

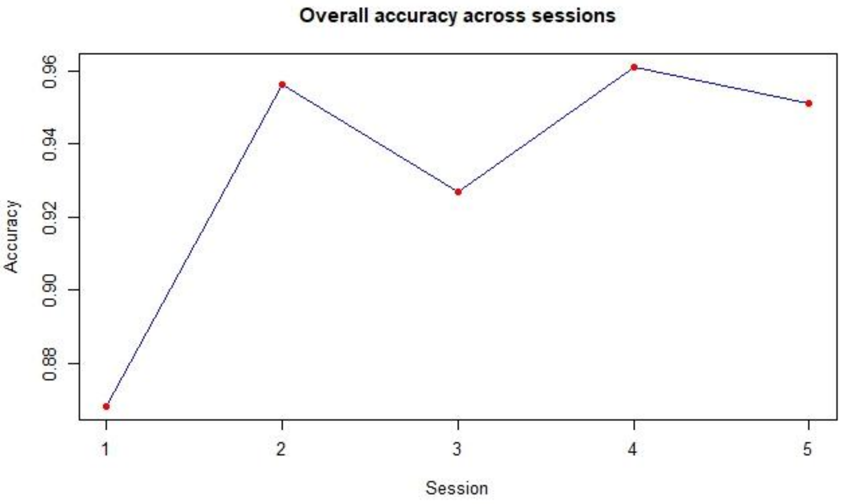


Figure 27: CSM overall accuracy across Sessions

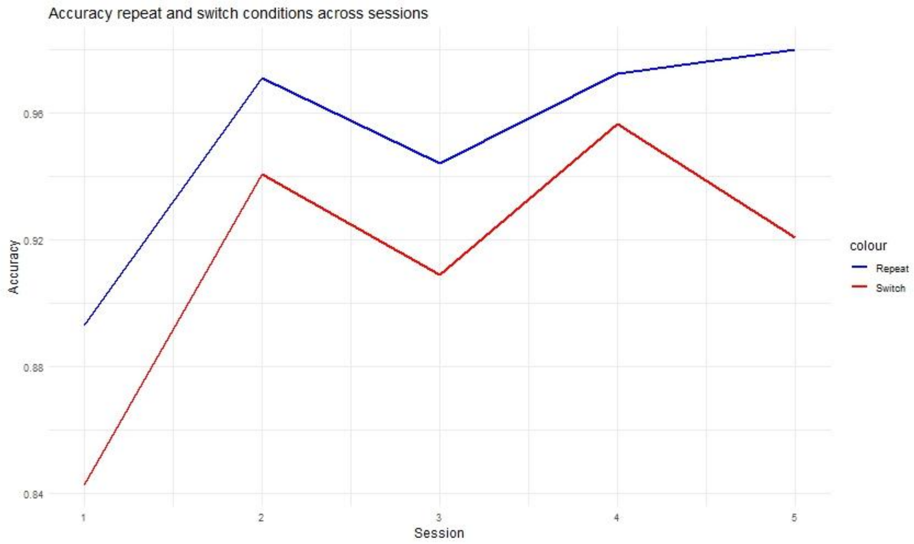


Figure 28: Progression of accuracy across sessions for switch and repeat conditions

By fitting a logistic regression model with Accuracy as the outcome variable and Session and Condition as predictor variables, the effect on Accuracy was analysed (see Figure 29).

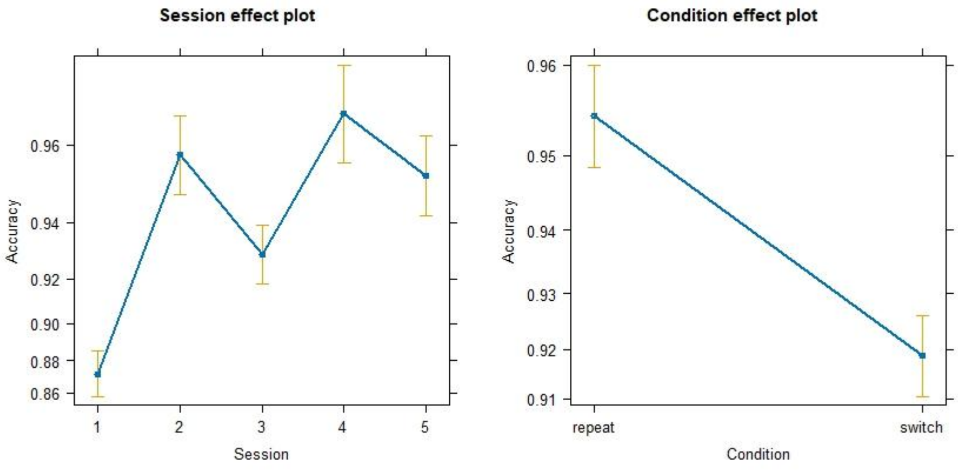


Figure 29: Linear mixed model for Accuracy across sessions

Results revealed highly significant effect 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 effect of Condition on Accuracy was also small ($f^2 = 0.03$) but highly significant ($p < 0.001$; SE = 0.08301), showing 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. Adding an interaction term to see whether Condition is modulated by Session, the model only revealed a significant interaction for Session 5 ($p < 0.01$, SE = 0.29822) with the switch condition showing a decreased accuracy compared to the repeat condition (see Figure 30). The overall effect size of the Session:Condition interaction was small ($f^2 = 0.03$).

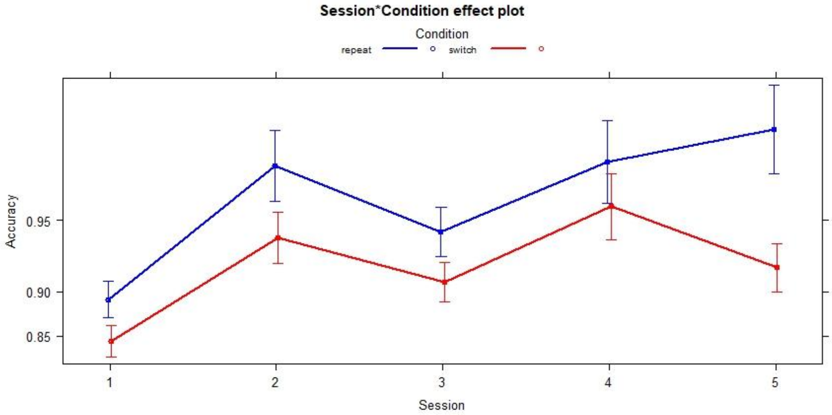


Figure 30: Interactions of Session and Condition across Sessions

The results that participants significantly decreased their overall RT across sessions until the end is consistent with the results by Dong & Liu (2016) and Babcock et al. (2017), who both used the colour shape switch task and found that participants had become faster after their respective training. However, as already mentioned in the Flanker task section, this effect occurred for all participant groups and was therefore not SI specific.

Although Figure 25 shows a numerical decrease of switch costs until Session 4, the linear mixed model did not reveal a significant effect of Session. This contradicts the findings by Dong & Liu (2016), who found that the interpreting groups showed significant gains in switch costs, therefore decreasing their switch costs significantly. The present finding that accuracy increased across testing contradicts the results by Dong & Liu (2016) and Babcock et al. (2017), who did not find an increase in accuracy at the post-test.

8.2.2.3 2-back

Baseline between-group comparison

Mean reaction time for the Control group was slightly lower than for the Control group with 0.581235ms (SD = 0.2602899) and 0.6024996ms (SD = 0.2271982) for the Experimental group. To test whether there is a significant effect of Group on reaction times, a linear mixed model with reaction time as outcome variable, Group as predictor variable and Participant as random effect was fitted (see Figure 31). Since the 2-back task required the participants only to press a button when they encountered a matching trial, only the match condition was included. Results revealed no significant difference in reaction time between groups (p = 0.801, SE = 0.04487) in the match condition (see Figure 31). However, as confirmed by a Q-Q plot and a Shapiro-Wilk test, residuals are not normally distributed. After log transforming the data to approach normal distribution, the results of the linear model did not differ from the original linear mixed model.

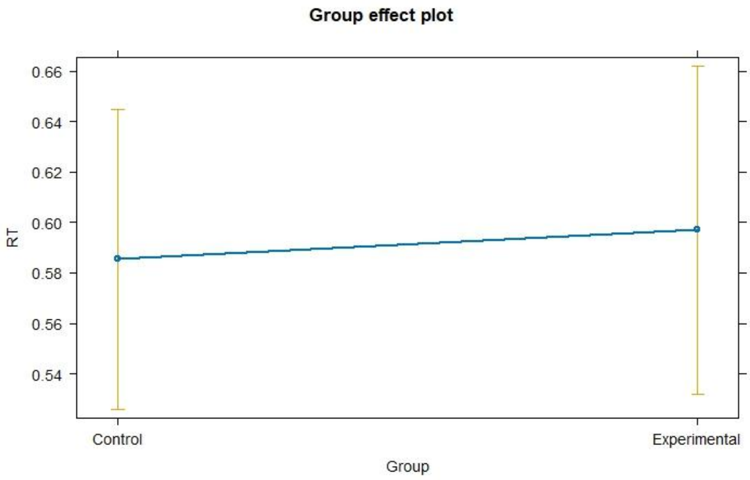


Figure 31: Linear mixed model for between-group differences in reaction times 2-back task

Table 17 includes mean reaction times of the match condition for both groups as well as the accuracy of match and the no-match condition.

Table 17: Descriptive statistics for Experimental and Control group

	Experimental group	Control group
Mean match RT (in ms)	0.602, SD = 0.227	0.581, 0.260
Overall accuracy	0.930, SD =0.256	0.889, SD = 0.314
Accuracy match	0.862, SD = 0.345	0.782, SD = 0.413
Accuracy no-match	0.961, SD = 0.193	0.940, SD = 0.238

A logistic regression model with Accuracy as outcome variable and Group and Condition as predictors was fitted to account for possible differences in accuracy between groups (see Figure 32). Results revealed a small (f^2 = 0.02) but significant effect of Group (p <0.001, SE = 0.09387) with an increase of the coefficient for the Experimental group by 0.52065 units, indicating a significantly higher accuracy for the Experimental group at baseline compared to the Control group. Furthermore, the effect of Condition was medium-sized (f^2 = 0.16) and significant (p < 0.001, SE =0.09166) with an increase by 1.43631 units for the no-match coefficient, indicating significantly higher accuracy for the no-match condition. There was no significant interaction between Condition and Group (p = 0.664; SE = 0.19106). VIF-values for both Condition and Group were equal to 1, not indicating any collinearity.

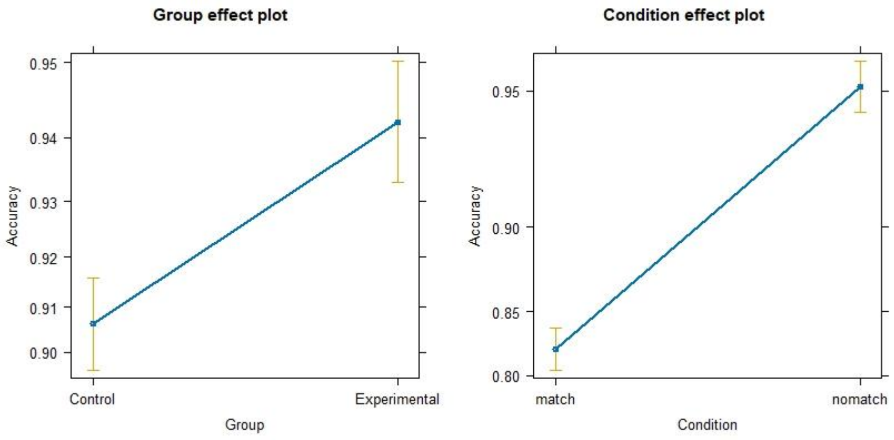


Figure 32: Logistic regression model for between-group comparison of accuracy in 2-back task

The results that there is no significant difference in RT between groups at baseline for the 2-back task correspond to results by Babcock et al. (2017) who did not find a difference between groups at baseline in two working memory measures and to results by Rosiers et al. (2019) and Dong & Liu (2016) who did not find a difference in RT between groups for the 2-back task at baseline testing. The result that accuracy was significantly higher for the Experimental group than for the control group at baseline contradicts the results by Rosiers et al. (2019) who did not find a difference between groups at baseline for accuracy in the 2-back task. The finding that the no-match condition was characterized by significantly higher accuracy corresponds to relevant 2-back literature such as by van der Linden et al. (2018) and Szmalec et al. (2011).

<u>Longitudinal results of Experimental group</u>

For the progression analysis of the Experimental group, Figure 33 depicts boxplots of the reaction times in the match condition across all five sessions. The mean reaction time seems to decrease over the course of testing.

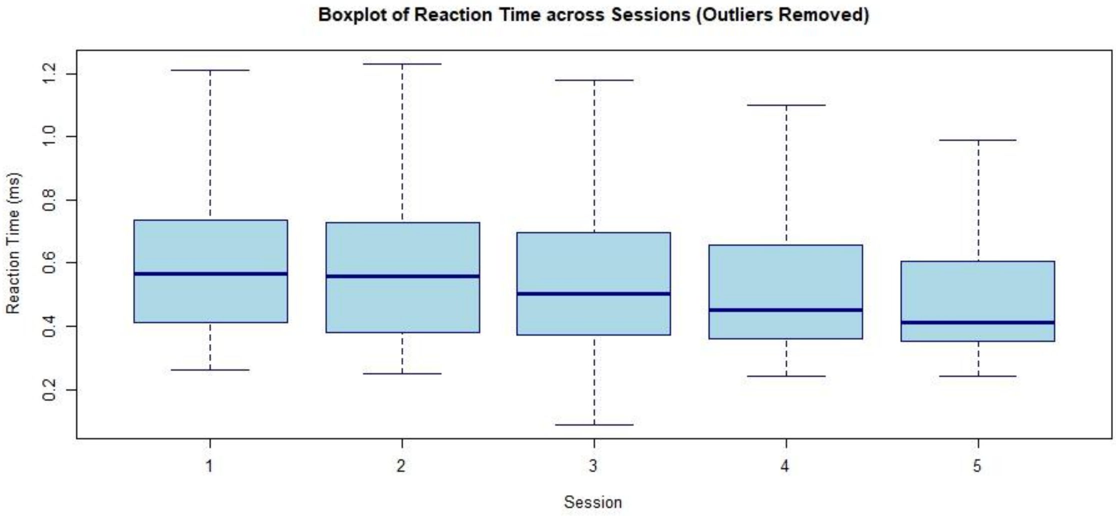


Figure 33: Boxplot of reaction times of match condition across sessions in 2-back task

To statistically confirm this observation, a linear mixed model was fitted with reaction time as outcome variable, Session as predictor and Participant as random effect (see Figure 34). Although results showed no significant effect of Session 2 on reaction time (p = 0.808033, SE = 1.078e-02), Session 3, Session 4 and Session 5 showed highly significant effects (S3: p = 0.000825, SE = 1.074e-02; S4: p = 0.001671, SE = 1.207e-02; S5: p < 0.001, SE = 1.145e-02) along with negative coefficients for all sessions. A

Q-Q plot of the residuals showed that the residuals are not normally distributed. Therefore, a log transformed linear mixed model was fitted. The results of the log transformed model did not differ greatly from the results of the previous linear mixed model.

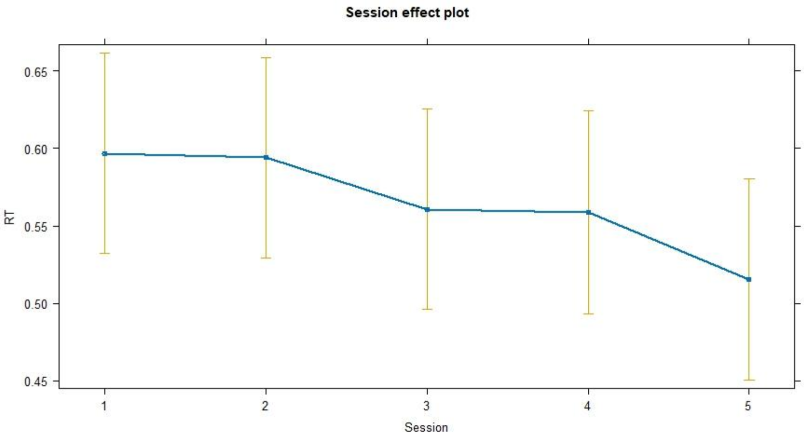


Figure 34: Linear mixed model for experimental group reaction times across sessions in 2-back task

Overall accuracy increased over the course of testing, as depicted in Figure 35. Taking a more differentiated look at accuracy in Figure 36 shows that accuracy in the match condition initially even decreases for Session 2 and then increased drastically, while accuracy for the no-match condition increased consistently and less drastically.

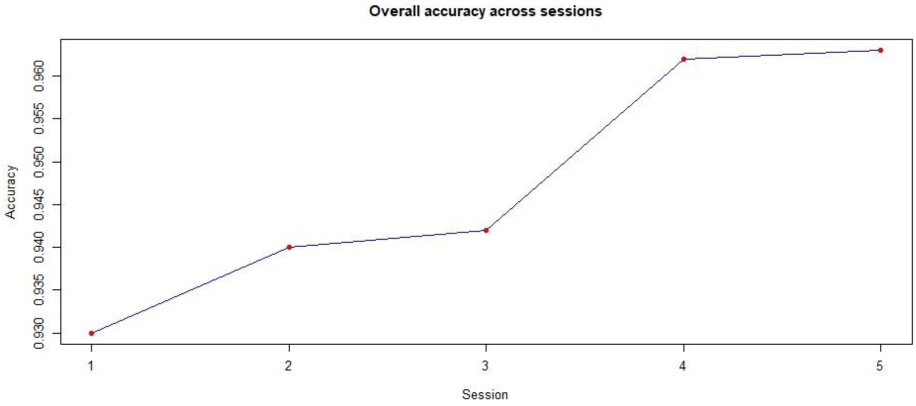


Figure 35: Overall accuracy across sessions for 2-back task

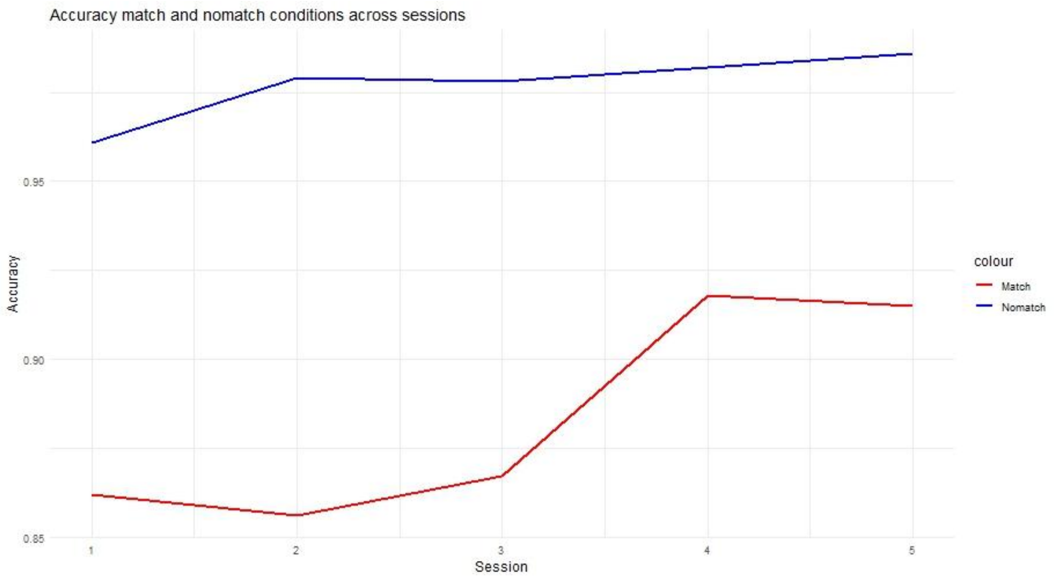


Figure 36: Accuracy for match and no-match condition across sessions for 2-back task

A logistic regression model with Accuracy as outcome variable and Session and Condition as predictors showed that effect of Session 2 was not significant (S2: $p = 0.1223$, SE = 0.11075), the effect of Session 3 was marginally significant ($p = 0.0517$, SE = 0.11188) and the effects of Session 4 and 5 were highly significant (S4: $p < 0.001$; SE = 0.14209; S5: $p < 0.001$, SE = 0.13633). The overall effect size of Session was small ($f^2 = 0.02$). The effect of Condition on accuracy was medium sized ($f^2 = 0.2$) and highly significant ($p < 0.001$, SE = 0.08612) with the no-match condition increasing by 1.73951 units compared to the match condition, showing that the match condition has significantly lower accuracy. VIF-values for Session and Condition were both equal to 1, excluding the possibility of multicollinearity.

Adding an interaction term to check for interactions between Session and Condition (see Figure 37) revealed significant interactions for Session 2 and Session 3 (S2: $p = 0.0042$, SE = 0.24068; S3: $p = 0.0288$, SE = 0.23835). The effect size of the Session:Condition interaction was small ($f^2 = 0.01$).

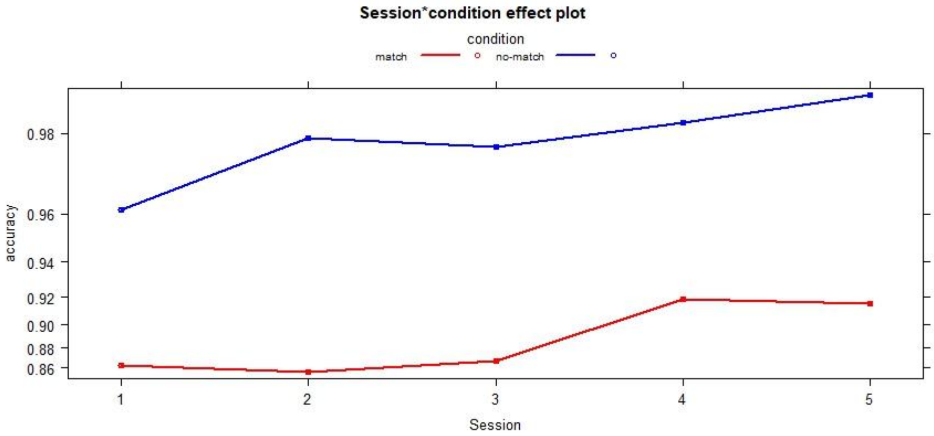


Figure 37: Interactions between Condition and Session for accuracy

The result that participants in the SI group significantly decreased their reaction times across sessions and became faster until Session 5 corresponds to the results by Dong & Liu (2016) who found that the SI group and not the bilingual control group improved their reaction times after training. The result that accuracy also increased over the course of testing, especially the significant increase in both condition in Session 4 and 5 corresponds to findings by Babcock et al. (2017) who found a numerical decrease of errors for the SI group and the control group.

8.2.2.4 3-back

Baseline between-group comparison

In the 3-back task, the linear mixed model to compare the two groups in reaction times at baseline level in the match condition did not reveal a significant effect of Group on reaction time (p = 0.177, SE = 0.05697), although the coefficient of the Experimental group did show an increase by 0.07981 units compared to the Control group (see Figure 38). A Shapiro-Wilk normality test revealed that residuals deviated from normality (p < 0.001; W = 0.96573). After fitting a log transformed linear model to normalize data distribution, results did not differ substantially from the previous linear mixed model.

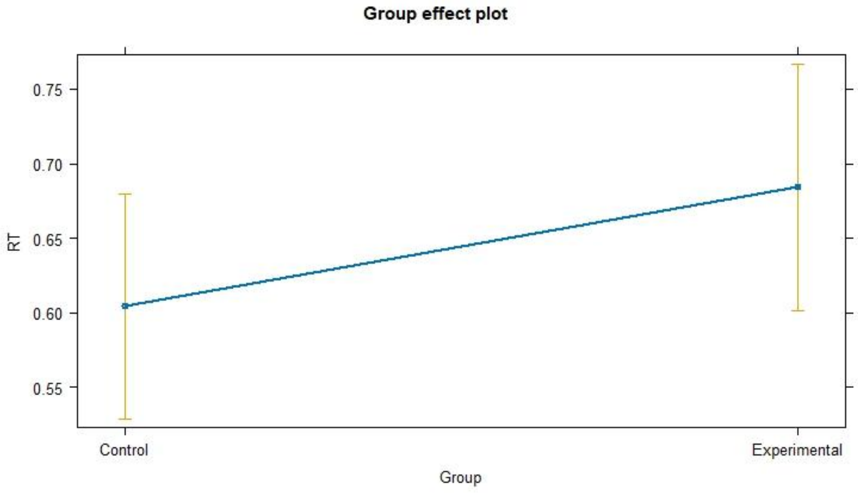


Figure 38: Linear mixed model of between-group comparison of reaction time at baseline for 3-back task

Mean reaction times per condition as well as the accuracy for both groups and conditions at baseline are depicted in Table 18. A binominal linear model with Accuracy as outcome variable, Group and Condition as predictor variables and Participant as random effect (see Figure 39) reveals no effect of Group on Accuracy (p = 0.41, SE = 0.07982) and a highly large (f^2 = 0.5) and significant effect of Condition on accuracy with an increase of the no-match coefficient by 2.76352 units, showing that significantly more errors were made in the match condition (p < 0.001, SE = 0.09079). There was no indication for collinearity between the predictors as revealed by the VIF-values that were equal to 1 for both Group and Condition. No significant interaction effect could be observed (p = 0.164, SE =0.185).

Table 18: Descriptive statistics for Experimental and Control group in 3-back task

	Experimental group	Control group
Mean match RT (in ms)	0.686, SD = 0.290	0.588, SD = 0.259
Overall accuracy	0.854, SD = 0.353	0.848, SD = 0.359
Accuracy match	0.608, SD = 0.488	0.608, SD = 0.488
Accuracy no-match	0.966, SD =0.181	0.957, SD =0.204

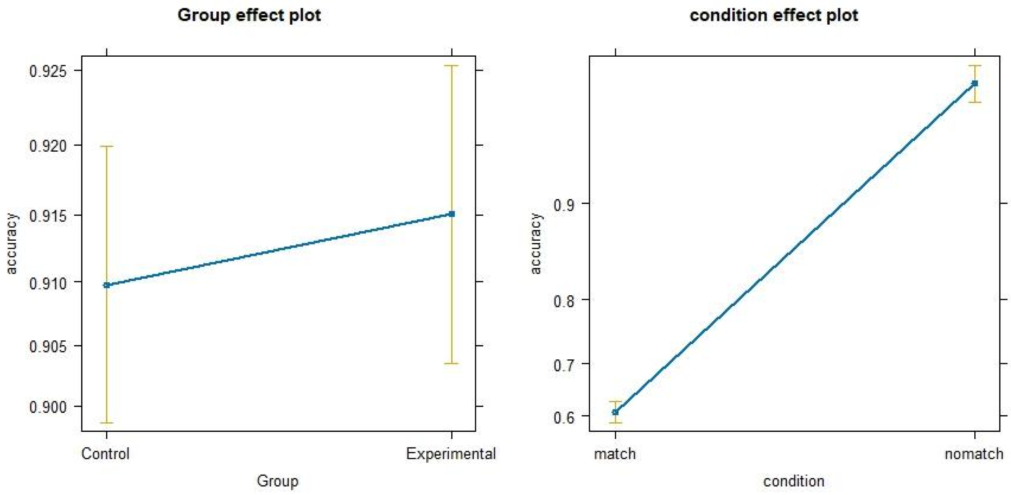


Figure 39: Between-group comparison of accuracy in the 3-back task

The findings that there was no significant group-difference in reaction times and accuracy at baseline for the 3-back task correspond to the findings by Dong & Liu and Rosiers et al. (2019) who used a 2-back task but did not find a significant difference between groups at baseline and to the findings by Babcock et al. (2017) who did not find a difference between groups in two different working memory measures. The findings that the no-match condition showed significantly higher accuracy than the match condition corresponds to relevant n-back task literature such as by van der Linden et al. (2018) and Szmalec et al. (2011).

<u>Longitudinal results of Experimental group</u>

Figure 40 shows the progression of mean reaction time across sessions, where it becomes apparent that mean reaction times decrease over the course of interpreting training and reaction times become shorter overall.

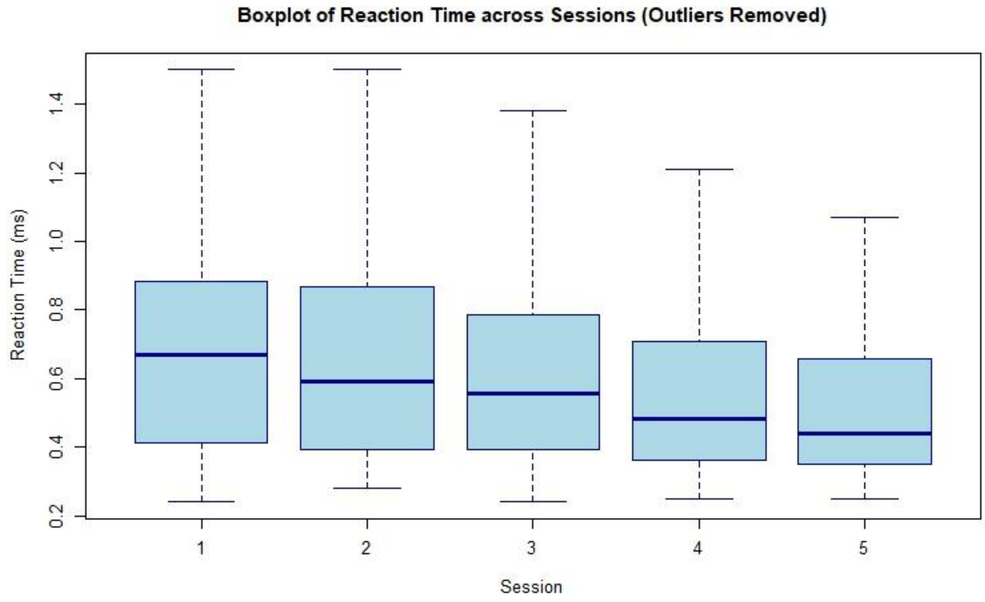


Figure 40: Progression of reaction time in match condition across sessions

To test the effect of Session on reaction time, a linear mixed model for the match condition was fitted with reaction time as outcome variable, Session as predictor and Participant as random effect. Results revealed significant effects of all Sessions but Session 2 (S2: p = 0.186, SE = 0.016; S3: p = 0.000192, SE = 0.01590; S4: p < 0.001, SE = 0.01782; S5: p < 0.0001, SE = 0.01693) and all Sessions did show negative coefficients (see Figure 41). Residuals revealed a non-normal distribution, which is why a log transformed linear model was fitted. Overall, log transformed results did not differ from the results of the original linear mixed model, only Session 2 showed a p-value that is now approaching significance (p = 0.069, SE = 0.02431). Results for the other log transformed Sessions continued to show very high significance as in the original model.

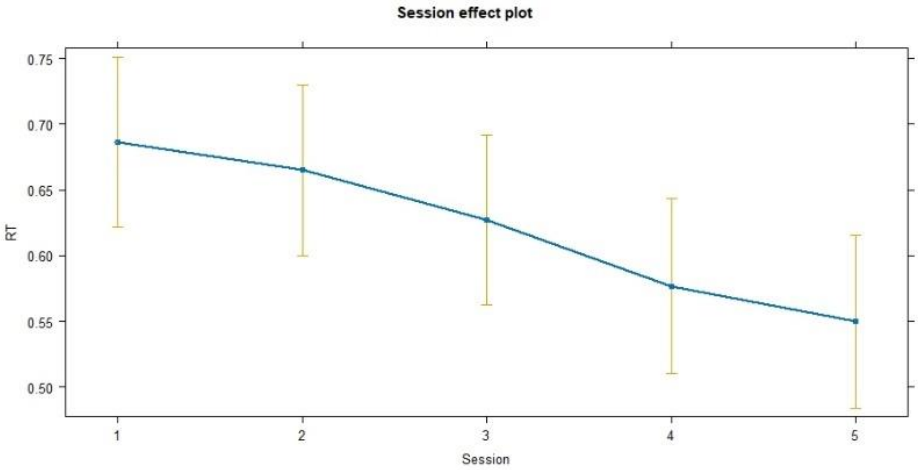


Figure 41: Linear mixed model for progression of reaction time across sessions for the 3-back task

As can be seen in Figure 42, overall accuracy increased steadily until Session 4, after which it decreased slightly. However, it becomes apparent in Figure 43 that accuracy in the match condition is not much lower in the 3-back task than in the 2-back task and furthermore, that while accuracy in the match condition increases consistently until the end of testing, accuracy in the no-match condition is more unstable.

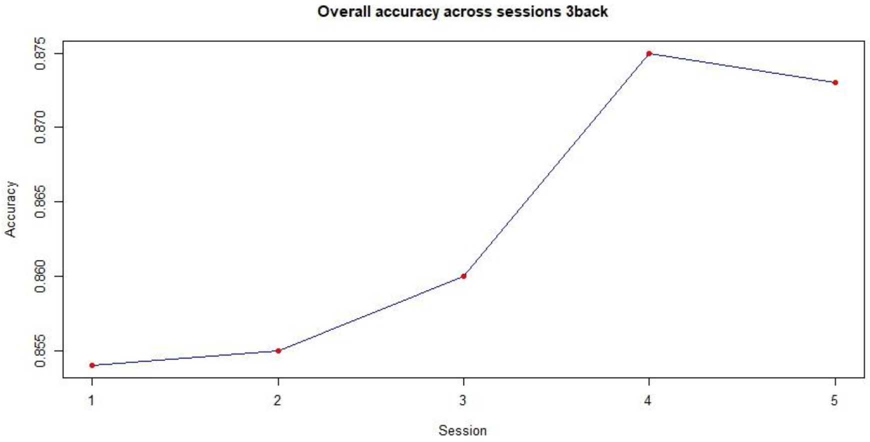


Figure 42: Overall accuracy in the 3-back task across sessions

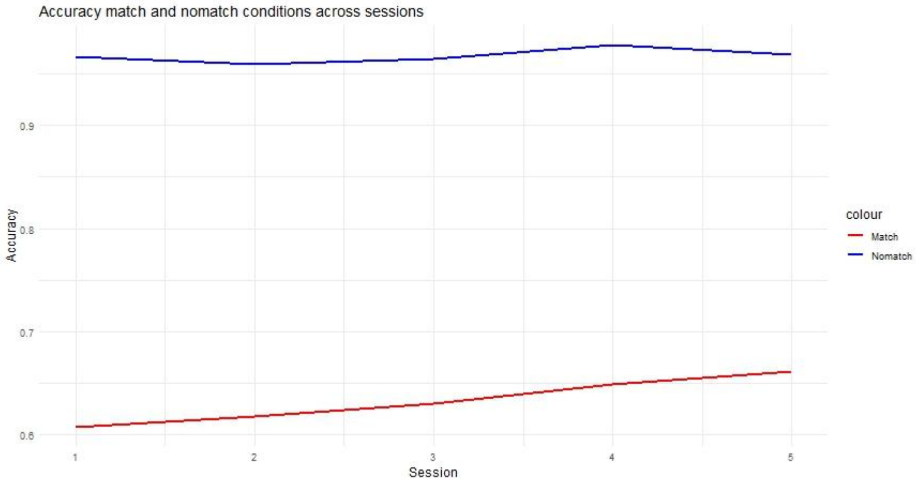


Figure 43: Accuracy in the 3-back task in both conditions across sessions

To test the effect of Session and condition on Accuracy, a logistic regression model with Accuracy as outcome variable and Session and Condition as predictors was fitted. Results showed no significant effects for Session 2 and 3 (S2: p = 0.8247, SE = 0.08430; S3: p = 0.5150, SE = 0.08468) but significant effects for Session 4 and 5 on Accuracy (S4: p = 0.0181, SE = 0.09555; S5: p = 0.0281, SE = 0.09152). The overall effect size of Session was very small ($f^2 < 0.01$). Finally, Condition had a significant effect on Accuracy (p < 0.001, SE = 0.06733) that was large in size ($f^2 = 0.52$), with an increase of the coefficient of the no-match condition by 2.82708 units, showing that the match condition had significantly lower accuracy compared to the no-match condition. There was no indication of collinearity between predictors as confirmed by VIF-values equal to 1 for both Session and Condition. Interactions between Session and Condition were not significant (all p > 0.05).

The result that participants in the SI group significantly decreased their reaction times across sessions and became faster until Session 5 corresponds to the results by Dong & Liu (2016) who found that the SI group and not the bilingual control group improved their reaction times after training. The result that accuracy also increased over the course of testing, especially the significant increase in both conditions in Session 4 and 5 corresponds to findings by Babcock et al. (2017) who found a numerical decrease of errors for the SI group and the control group. The highly significant effect of Condition on Accuracy with the match condition showing lower accuracy corresponds to relevant n-back task literature (see van der Linden et al. (2018) and Szmalec et al. (2011)).

8.2.2.5 2-back vs. 3-back task

Lastly, reaction times and accuracy across sessions were compared between the 2-back and the 3-back task for the interpreting group. Initially, a linear mixed model with reaction times as dependent variable and Session and Task (2-back/3-back) and Participant as random effect was fitted. Data was not normally distributed, which is why a log-transformed linear mixed model with the same variables was fitted and used for calculations (see Figure 44). Results showed a significant main effect of Session, with all Sessions showing significant effects on the reaction time except Session 2, which only displayed a marginally significant effect on reaction time (S2: $p =0.055$; SE = 0.01444; S3: $p < 0.001$, SE = 0.01436; S4: $p < 0.001$, SE = 0.01596; S5: $p > 0.001$; SE = 0.01523). The overall effect size of Session was small as indicated by Cohen's f^2 ($f^2 = 0.02$). Furthermore, there was a small ($f^2 = 0.02$) but significant main effect of task ($p < 0.001$; SE = 0.01082) with the coefficient of the 3-back task increasing by 0.10508 units, showing slower reaction times for the 3-back task than for the 2-back task. VIF-values for both Session and Task were equal to 1 and did not indicate any collinearity.

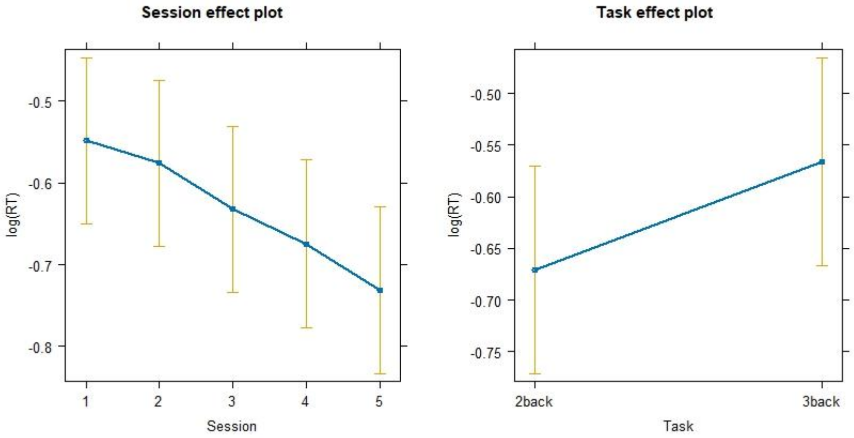


Figure 44: Linear mixed model of the reaction times of the 2-back and the 3-back task

Adding an interaction term revealed a small ($p < 0.01$) but significant interaction for Session 4 ($p = 0.014$, SE = 0.03216) where the 3-back task seemed to decrease more compared to the 2-back task (see Figure 45).

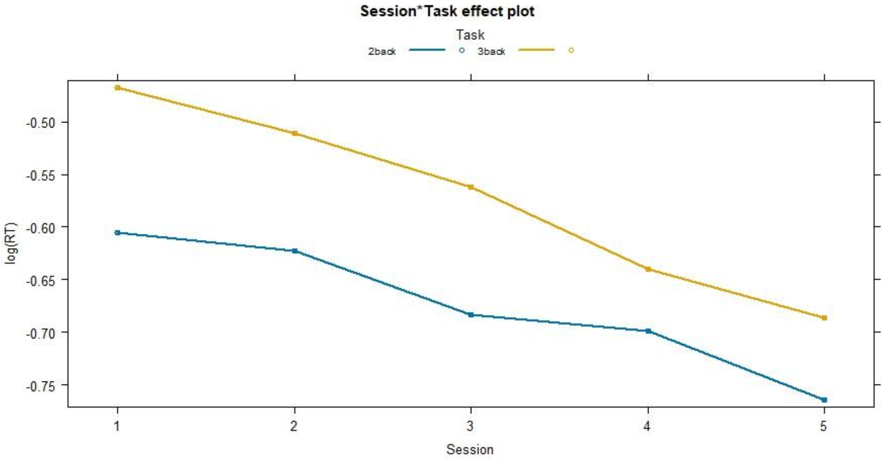


Figure 45: Interaction model of reaction times in 2-back and 3-back tasks

In terms of accuracy, a logistic regression model was fitted with Accuracy as dependent variable and Task and Session as independent variables (see Figure 46). Results revealed a medium-sized ($f^2 = 0.16$) and highly significant main effect of Task ($p < 0.001$; SE = 0.058496) with the 3-back task showing significantly lower accuracy than the 2-back task. In terms of the main effect of Session on accuracy, only Session 4 and Session 5 did show significant effects (S2: $p = 0.9346$; SE = 0.079485; S3: $p = 0.3307$, SE = 0.079805; S4: $p = 0.0015$; SE = 0.090808; S5: $p < 0.001$, SE = 0.087440) and the overall effect size of Session was very small ($f^2 < 0.01$). The VIF-values for Session and Task were both equal to 1, not indicating any collinearity between predictors.

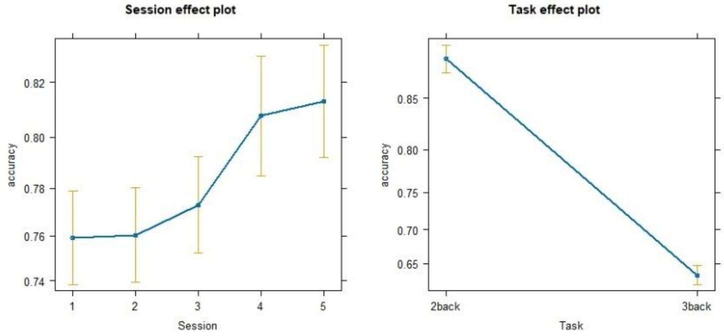


Figure 46: Logistic regression model of accuracy in 2-back and 3-back task across sessions

Adding an interaction term revealed only a significant interaction of Session and Accuracy in Session 4 ($p = 0.04600$, $SE = 0.20680$) with the 2-back task increasing more in Session 4 than the 3-back task (see Figure 47). The effect size of the Session:Task interaction term was small $f^2 = 0.01$.

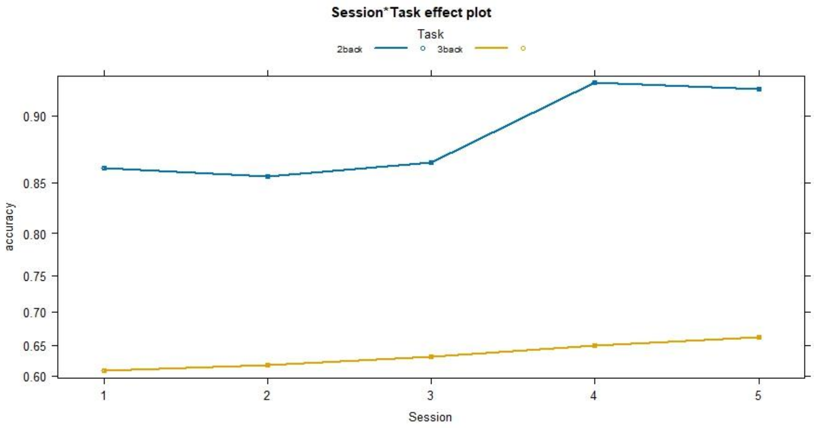


Figure 47: Interaction model of accuracy in 2-back and 3-back comparison across sessions

The result of longer reaction times and lower accuracy for the 3-back than for the 2-back task corresponds to relevant n-back literature (Jaeggi et al. 2010) explaining that processing load can be varied with increasing N, so 3-back requires more cognitive processing than the 2-back task, more specifically different processes of inhibition and working memory updating. Furthermore, the results correspond to findings such as by Morales et al. (2015) that found that 2-back task was easier to perform than the 3-back task.

8.2.3 ERP data

For the Flanker task, the ERP component N200 that has been related to inhibition and monitoring processes (see Section 8.1.2) was examined. A fronto-central AOI of the electrodes f1, fz, fc1, fcz was selected and analysed within the time window of 200-400ms based on literature listed in Section 8.1.2.

The focus of the colour-shape switch task lay on the ERP components P300 and N200. The P300 component has been related to mental flexibility, working memory

updating and monitoring (see Section 8.1.2). Based on literature listed in Chapter 8.1.2, a fronto-parietal AOI of the electrodes cz, cp1, cpz and p1 was selected and analysed within the P300 time-window 280-400ms. For the N200 time window 200-320ms, a frontal AOI of the electrodes fz, fcz and cz was selected.
For the 2-back and 3-back tasks, the P300 ERP component was examined.

8.2.3.1 Flanker task

Initially, voltages of the Experimental group and Control group were compared at baseline within the respective N200 time window by fitting a linear mixed model with Voltage as outcome variable, Group and Condition as predictor variables and Participant as random effect (see Figure 48). Data was normally distributed as indicated by a Shapiro-Wilk test (p = 0.06318, W = 0.97356). Results indicated no significant effect of Group on Voltage (p = 0.995, SE = 1.176029). This result corresponds to previous studies that did not find a behavioural difference between groups at baseline in tasks that target the inhibitory executive functions (see Dong & Liu (2016), Babcock et al. (2017), Rosiers et al. (2019)). Furthermore, there was no significant effect of Congruency on voltage (p = 0.574, SE = 0.245804). Adding an interaction term to check for an interaction of Group and Condition did not reveal a significant effect (p = 0.647, SE = 1.35648). This result contradicts the hypothesis based on N200 literature that assumed more negative voltage for the incongruent condition in the N200 time-window (see Wild-Wall et al. (2008), Johnstone et al. (2009), Kousaie & Phillips (2012), Grundy et al. (2017b)). However, based on these results, the interpreter group did not show an initial advantage in inhibition or monitoring abilities over the Control group at baseline. VIF-values for both Condition and Group were equal to 1, excluding the possibility of multicollinearity.

Figure 48: Linear mixed model of between-group comparison of voltage at baseline

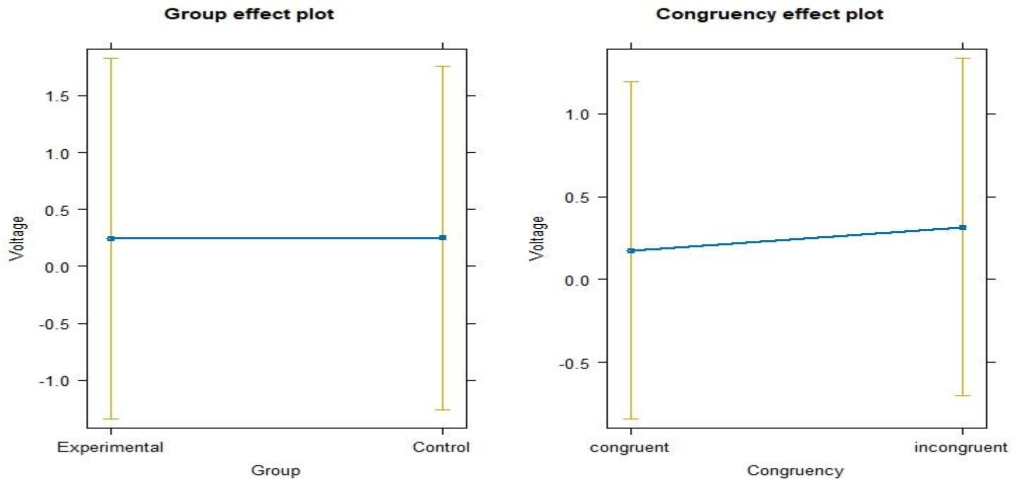


Figure 48: Linear mixed model of between-group comparison of voltage at baseline

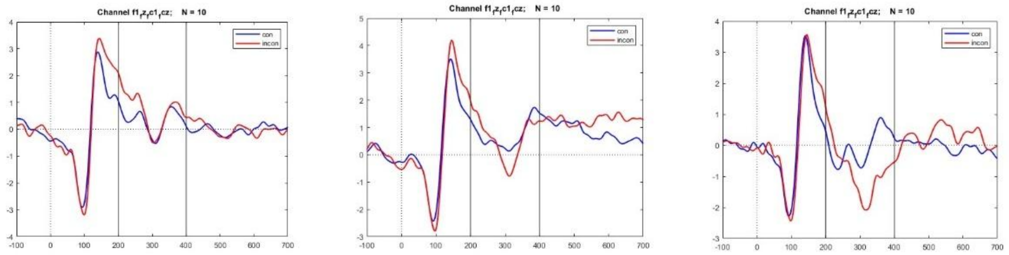
Table 19 shows mean voltage as baseline for both groups per condition as well as the Flanker interference values at baseline. Two more linear models were fitted to analyse the effect of the congruent and incongruent condition on Voltage separately. There was neither a significant effect of Group on Voltage in the congruent condition (p =0.790, SE = 1.12707), nor in the incongruent condition (p =0.772, SE =1.0888).

Table 19: Descriptive statistics for both groups and conditions at baseline for Flanker

	Experimental group	Control group
Mean voltage congruent	0.024009, SD =2.81738	0.3284109, SD = 2.344885
Mean voltage incongruent	0.4935449, SD = 2.267134	0.1698827, SD = 2.689779
Flanker interference	-0.4695359	0.1585282

Figure 49 shows the ERPs in the N200 time-window for the simultaneous interpreting group across all sessions. Amplitudes that resemble a relatively late N200 component can be detected.

Session 1 (Baseline) Session 2 Session 3



Session 4 Session 5

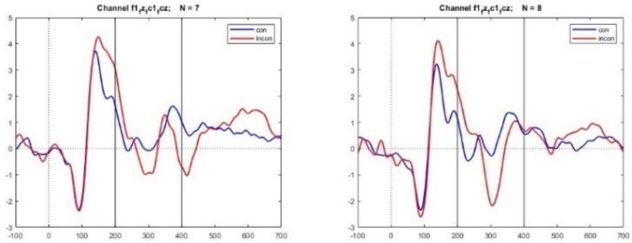


Figure 49: Flanker ERPs of interpreting group across sessions

In the form of a topographical 2D-map of the scalp, Figure 50 shows the distribution of the averaged voltage at the time window 200-400ms after stimulus onset on the scalp per session and per condition. Interesting for the inhibition process are especially the fronto-central electrodes that are associated with the executive function of inhibition and monitoring. As indicated by the blue colour, voltages in the frontal region become slightly more negative, especially in the incongruent condition.

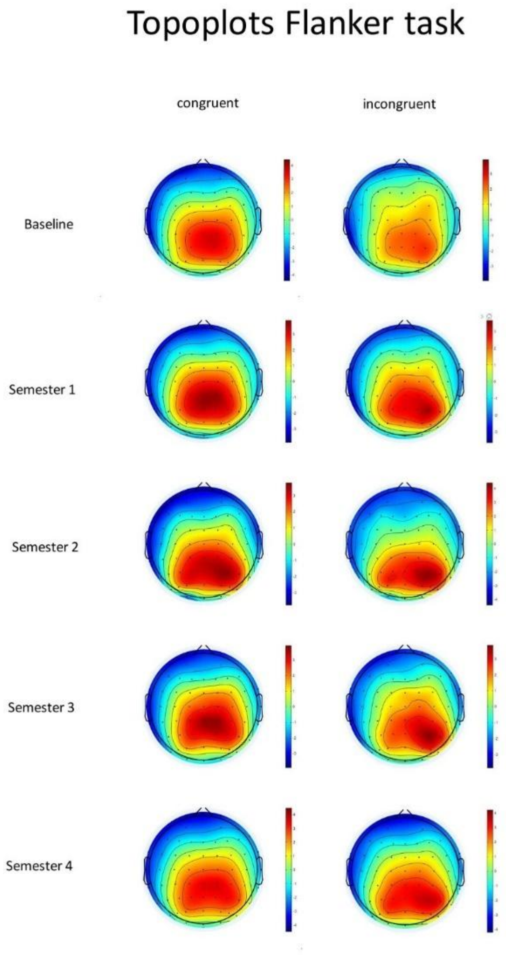


Figure 50: Topoplots Flanker task

The progression of mean voltage in the N200 time-window is depicted in Figure 51. Mean voltage especially in the incongruent condition seems to decrease until Session 3 where it hits a negative peak. Subsequently, mean voltage increases again for Session 4, followed by a slight decrease for Session 5.

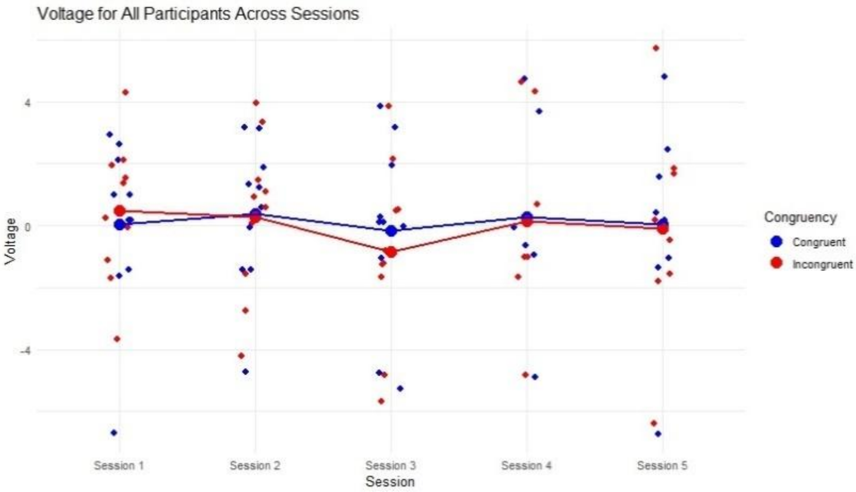


Figure 51: Progression of mean voltage within N200 time-window across sessions per condition

A linear mixed model with Voltage as outcome variable, Session and Condition as predictor variables and Participant as random effect was fitted to test the influence of Session and Condition on voltage. Data showed normal distribution as indicated by a Shapiro-Wilk test (p = 0.4419; W = 0.97389). The results of the model are depicted in the following Figure 52:

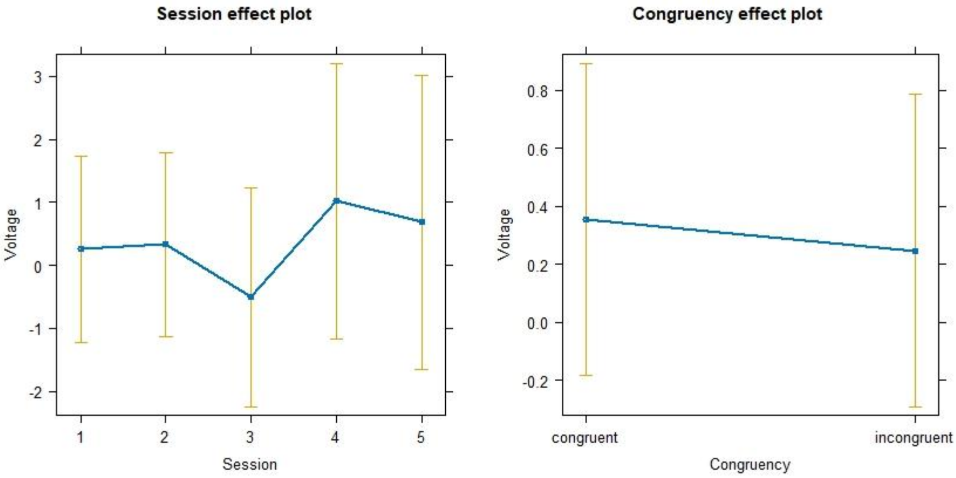


Figure 52: Linear mixed model of voltage across sessions in N200 time-window of Flanker task

Although the linear mixed model in general mirrors the trends of the mean voltage progression depicted in Figure 51, no effect of Session was significant. (S2: p= 0.857, SE = 0.38105; S3: p = 0.141, SE = 0.47626; S4: p = 0.675, SE = 1.75289; S5: p = 0.818; SE = 1.81166). The fixed effects coefficients show a slight increase by 0.07037 units for Session 2 to Baseline, a decrease by -0.76178 units for Session 3, an increase by 0.76086 units for Session 4 and an increase by 0.43168 units for Session 5.

Interestingly, the effect of Congruency was not significant either (p = 0.420, SE = 0.13220) although the fixed effects coefficient shows a decrease by -0.10741 units for the incongruent condition. The VIF values for both Session and condition were equal to 1, indicating that no collinearity is present.

Additionally, after adding an interaction term to see if Congruency is modulated by Session, results did only show a significant interaction of Session 3 and Congruency (p = 0.00272; SE = 0.36301). The effect seemed to be driven by the incongruent condition, showing that voltage decreased more drastically in the incongruent condition at Session 3 compared to the congruent condition (see Figure 53). The effect size of the Session:Condition interaction was small ($f^2 = 0.01$).

The lack of effect of Session on Voltage across the time of testing and therefore also the lack of evidence that participants improved their interference suppression abilities over SI training corresponds to van de Putte et al. (2018), who did not find a behavioural improvement in inhibitory control after 9 months of training.

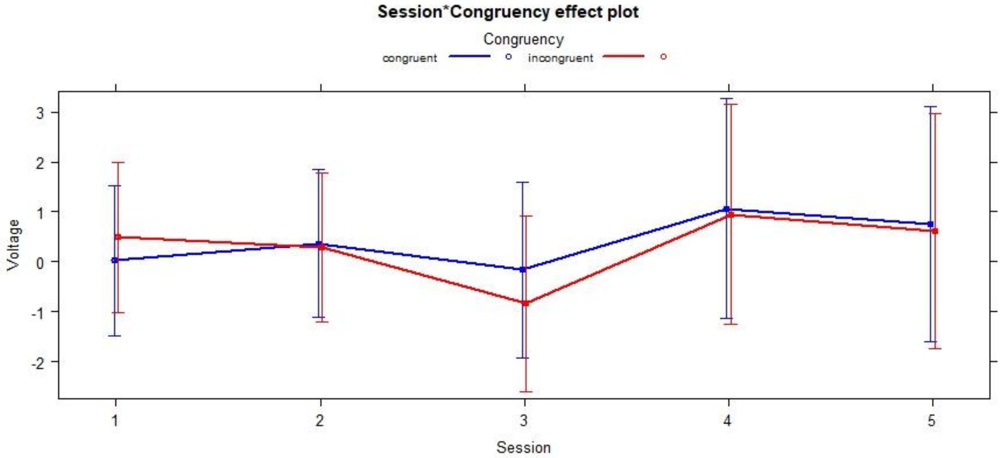


Figure 53: Interactions for Session and Congruency across Sessions

Concerning the electrophysiological Flanker interference effect, Figure 54 shows that the difference between condition increases until Session 3, after which it decreases again.

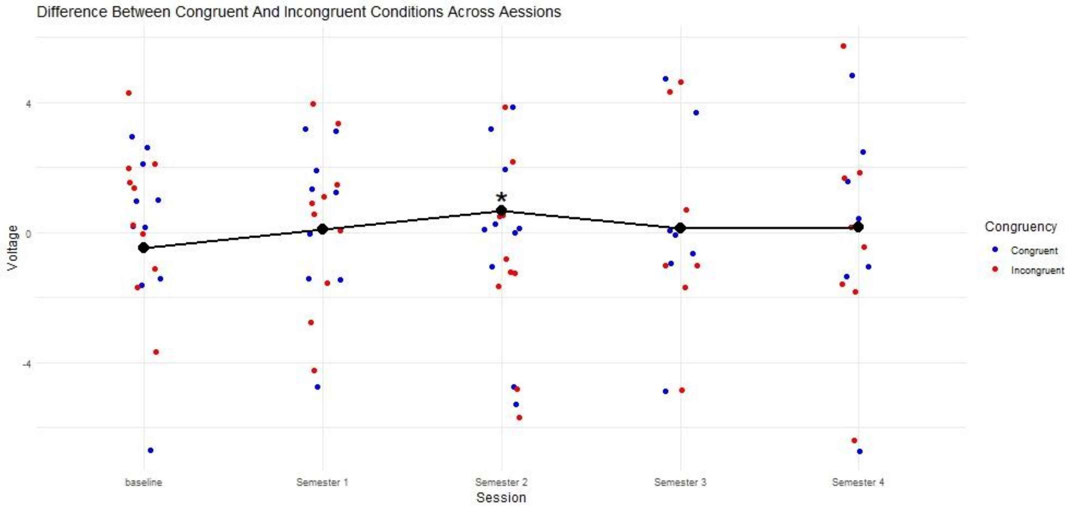


Figure 54: Electrophysiological Flanker interference effect across sessions

Finally, the peak latency within the N200 time-window in the Flanker task was analysed by computing the most negative peaks for both conditions in the N200-window. It becomes apparent in Figure 55 that the incongruent condition always shows longer latencies compared to the congruent condition. Furthermore, latencies increase steadily for the incongruent condition, while latencies for the congruent condition initially decrease drastically until Session 3 and then slightly increase until Session 4 which is followed by a plateau.

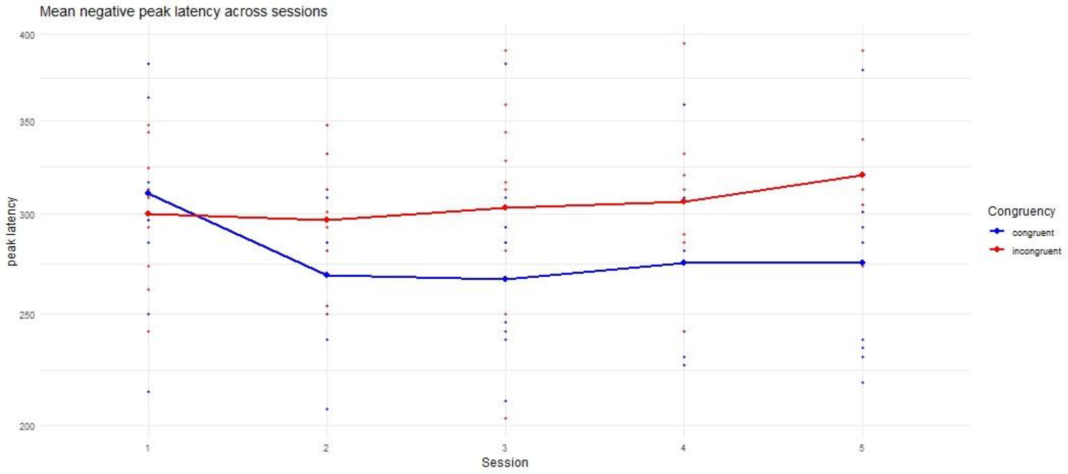


Figure 55: Peak latency Flanker in N200 time-window

A log transformed linear model revealed no significant effects of Session on latency (all Session p > 0.05) but a small- to medium-sized ($f^2 = 0.08$) and significant effect of Congruency (p = 0.005, SE = 0.031) with an increase of the incongruent condition by 0.08782 units, showing significantly higher latencies for the incongruent condition. This corresponds to findings in relevant ERP-literature that associate longer latencies with more processing effort which is given in the incongruent condition in comparison with congruent conditions (see Cespón & Carreiras (2020)).

VIF-values for Session and Condition were equal to 1 and did not reveal any sign of collinearity. Including an interaction term to check for an interaction between Congruency and Session did not reveal significant results (all p > 0.05).

To depict the effect that participant variability might have had on the results of the longitudinal analysis, Figure 56 presents the progressions of mean voltage per participant and condition. Although there seems to be a trend for a negative peak at Session three as it can be seen in Figure 51, Figure 52 and Figure 53, variability between participants is high.

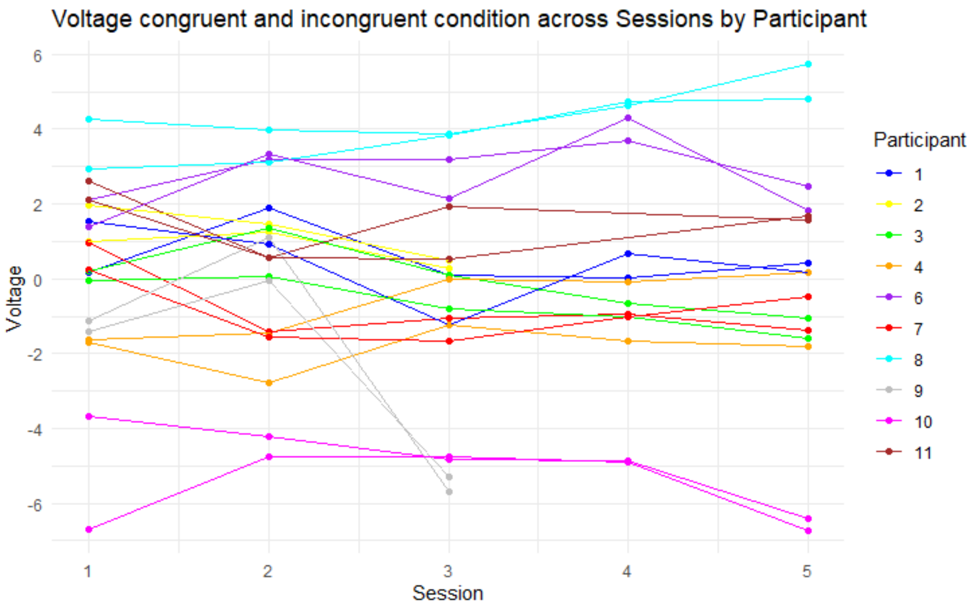


Figure 56: Mean voltage across Sessions for all participants and conditions

8.2.3.2 Colour-shape switch task

P300 time-window

For the P300 time-window, Table 20 depicts mean voltage in both conditions per group as well as the switching costs as baseline, showing on the one hand that the mean voltage in the P300 time-window was more negative for the switch condition than the repeat condition and that the mean voltage in the P300 time-window was lower for the control group than for the experimental group in both conditions.

Table 20: Mean voltage in switch and repeat condition per group for the colour-shape switch task in the P300 time-window

	Experimental group	Control group
Mean voltage switch P300	0.2285498, SD =1.3293	-0.6886222, SD=0.9682297
Mean voltage repeat P300	0.4580216, SD =1.314701	-0.06820417, SD=0.9137492
Switching costs P300	0.2294718	0.620418

Voltages within the P300 time-window of the Experimental group and the Control group were compared at baseline level by means of a linear mixed model with Voltage as outcome variable, Group and Condition as predictor variables and Participant as random effect. Data was normally distributed as indicated by a Shapiro-Wilk test (p =0.6214 ; W = 0.97969). Results are depicted in Figure 57 and indicated small- to medium sized ($f^2 = 0.1$) significant effect of Group (p =0.0393, SE = 0.3389) with Control group decreasing by -0.7217 units, indicating that the Control group showed significantly lower voltages than the Experimental group. Although the switch condition decreased by -0.4427 units, the effect of Condition was 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. An interaction term was fitted to see whether Condition is modulated by Group, however, there was no significant interaction (p = 0.571; SE = 0.6834).

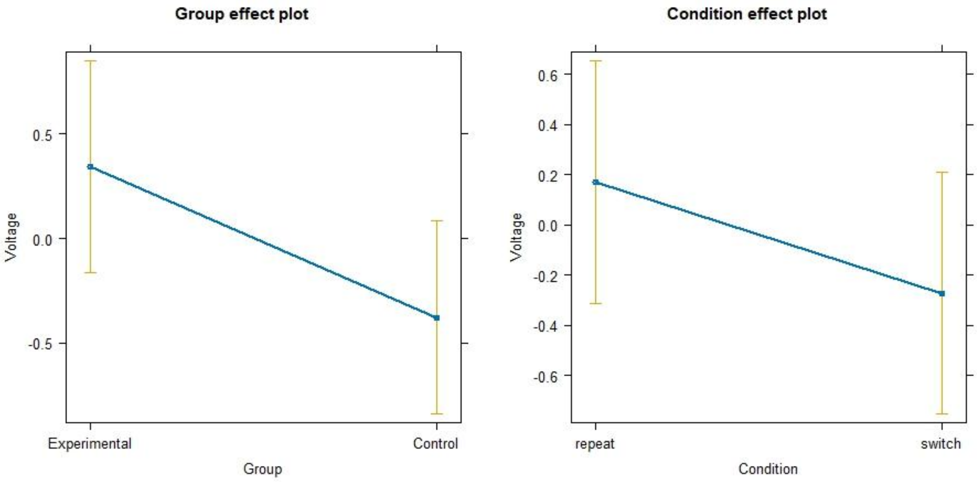
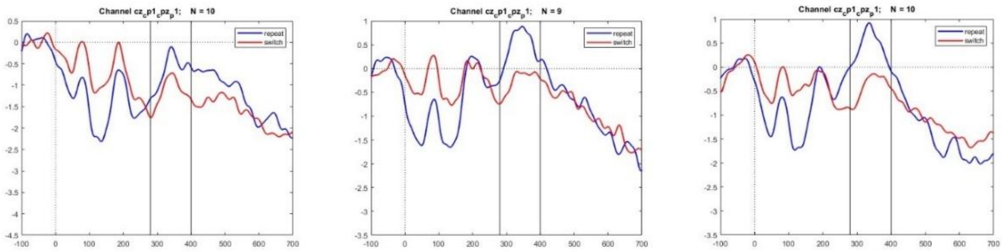


Figure 57: Linear mixed model P300 component Experimental vs. Control group

The finding that there was a significant effect of group with the control group showing significantly lower voltages than the Experimental group contradicts previous studies that did not find a behavioural difference between groups at baseline in the colour-shape switch task (Babcock et al. (2017), Dong & Liu (2016) and Rosiers et al. (2019)).

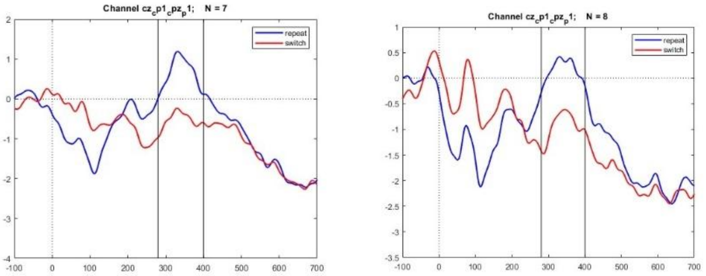
For the Simultaneous interpreters, amplitudes resembling the P300 component are depicted in Figure 58 for all Sessions:



Session 1 (Baseline)

Session 2

Session 3



Session 4

Session 5

Figure 58: P300 ERPs for CSM (all sessions)

Depicted in a 2D topographical map, Figure 59 shows the distribution of the averaged voltage at the time window 280-400ms after stimulus onset on the scalp per session and per condition. For the executive function of task switching as well as the P300 component distribution, central and parietal regions are relevant. By visual inspection it can be seen that voltage in these regions becomes more positive especially for the switch condition over the course of interpreting training.

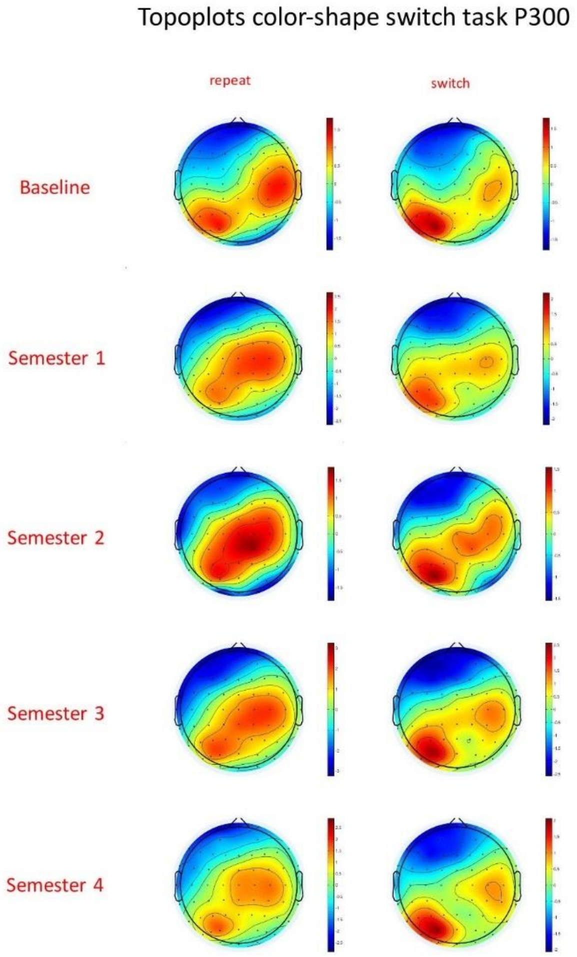


Figure 59: Topoplots colour-shape switch task P300 time-window

The progression of mean voltage for the Experimental group (see Figure 60) for switch and repeat conditions in the P300 time window shows a steady increase for both conditions from baseline and peak of mean voltage at Session 4, followed by a decline of mean voltage almost back to baseline.

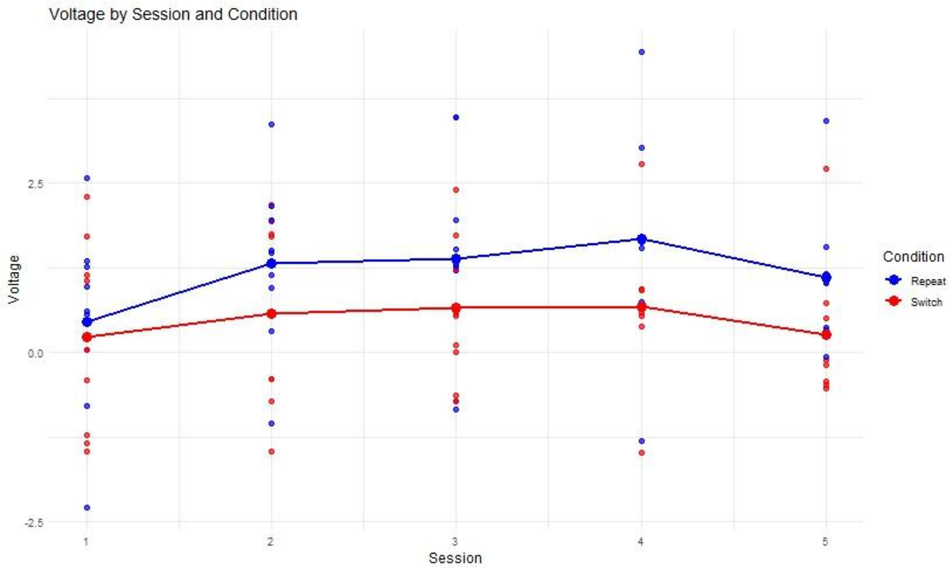


Figure 60: Progression of mean voltage in P300 time window

This trend is confirmed by a linear mixed model (see Figure 61) fitted with Voltage as dependent variable, Session and Condition as predictor variables and Participant as random effect. Data showed a normal distribution as indicated by a Shapiro-Wilk test (p = 0.1066; W = 0.9763). While the effect of Session 2 is not significant (p =0.08058, SE =0.3403), the effect of Session 3 and 4 on voltage are significant (S3: p =0.04076, SE =0.3295; S4: p =0.03796, SE = 0.3704). The effect of Session 5 on Voltage is not significant (p =0.45043, SE =0.3543), unsurprisingly so since voltage seems to return almost back to baseline. Overall, the variable of Session shows a small effect as indicated by the Cohen's f^2 ($f^2 = 0.05$). The effect of Condition is small ($f^2 = 0.07$) but highly significant (p =0.00308, SE = 0.2221) with switch condition decreasing by -0.6799 units, showing that the switch condition is significantly more negative than the repeat condition. The VIF for both Session was equal to 1, indicating that the independent variables do not show multicollinearity. Integrating an interaction term of Condition in the linear mixed model to examine how the effect of Session on Voltage changes depending on Condition yielded non-significant results for all sessions (all p>0.05).

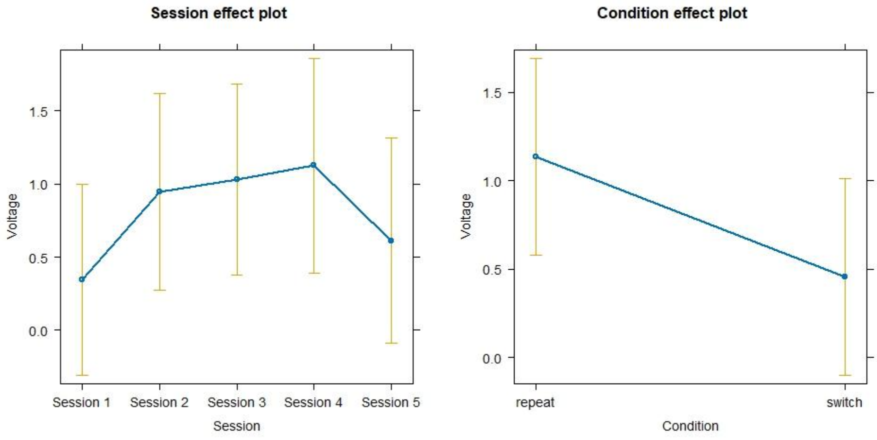


Figure 61: Linear mixed model P300 across sessions

The finding that the effect of Session 5 on voltage did not reveal a significant result contradicts the findings from the behavioural data as well as previous behavioural studies that found a improvement in task switching performance at the end of a longitudinal testing design (see Dong & Liu (2016) and Babcock et al. (2017)). The significant main effect of Condition on Voltage with the repeat condition showing more positive voltage than the switch condition corresponds to relevant P300 task switching literature (see e.g. López Zunini et al. (2019), Chen et al. (2022), Periáñez & Barceló (2009), Cespón & Carreiras (2020)).

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 which corresponds to the behavioural findings but contradicts findings from previous behavioural studies such as by Dong & Liu (2016). The progression of electrophysiological switching costs is depicted in Figure 62:

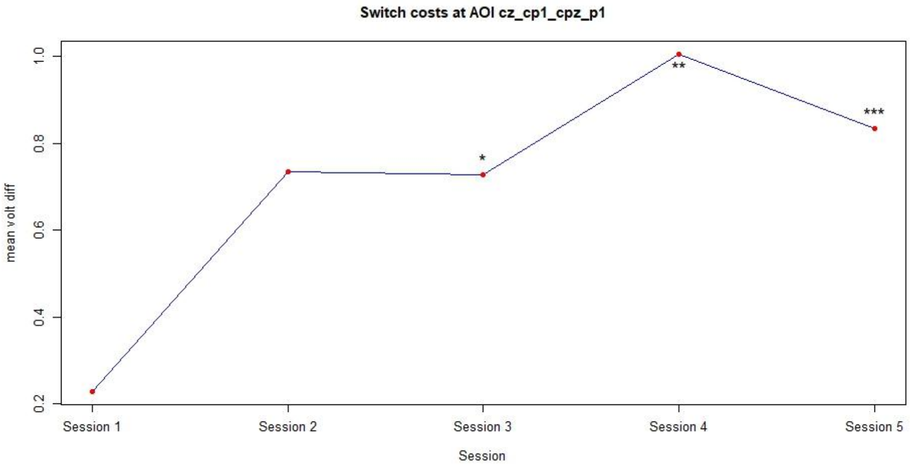


Figure 62: Electrophysiological switching costs P300 across sessions

To analyse peak latency for the colour-shape switch task, the most positive peaks within the P300 time-window were computed for both conditions across all sessions. Figure 63 shows great variability in latencies for both conditions. While latencies initially decreased for both conditions, latency of the repeat condition decreased further until Session 3, which is followed by an increase for Session 4 and another decrease for Session 5. After the initial decrease of the switch condition, latency increased again for Session 3 which is then followed by a sharp decrease for Session 4, the counter development of the repeat condition at this time-point. Latency of the switch condition then seems to remain steady for Session 5.

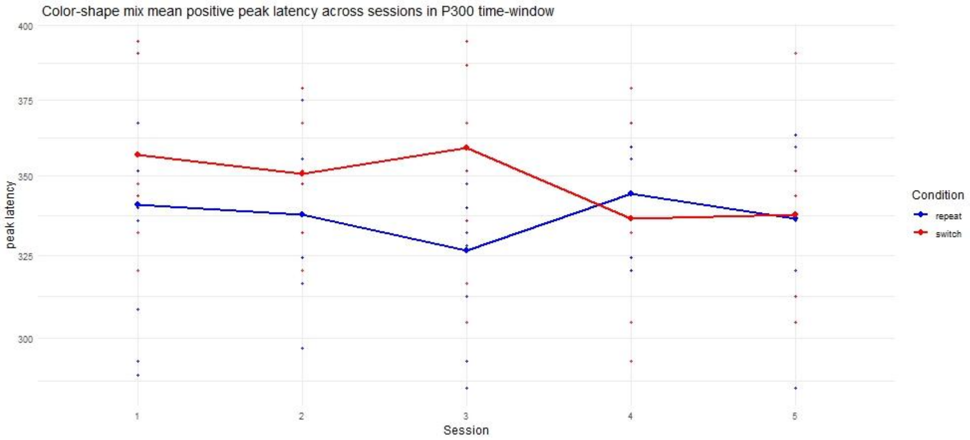


Figure 63: Peak latency for colour-shape switch task in P300 time-window

A linear mixed model was fitted to test the effect of Session and Condition on latency. Data was normally distributed. Although the regression coefficients indicate a decrease in latency across Sessions (see Figure 64), results revealed no significant effects of Session (all Session $p > 0.05$). The effect of Condition however was small ($f^2 = 0.05$) but significant ($p = 0.024$, SE = 5.492) with the switch condition regression coefficient increasing by 12.695 units, showing significantly longer latencies for the switch condition compared to the repeat condition. This corresponds to findings from relevant P300 task switching literature that associates switch trials with longer P300 latencies than repeat trials (see Céspon Carreiras (2020)). No collinearity was detected for the predictors Session and Condition indicated by VIF-values equal to 1. Interactions between Condition and Session did not show significant results (all $p > 0.05$).

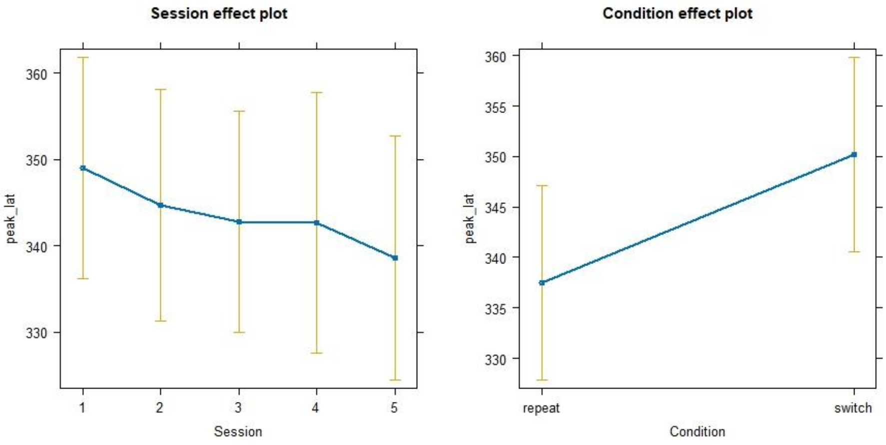


Figure 64: Linear mixed model of peak latency for colour shape switch task in P300 time-window

While no effect of Session on latency was found, the effect of Condition on Voltage was significant which corresponds to findings from relevant P300 task switching literature that associates switch trials with longer P300 latencies than repeat trials (see Céspon Carreiras (2020)).

N200 time-window

Voltages for the N200 time-window at baseline were compared between the Experimental and the Control group by means of a linear model with Voltage as outcome variable, Group and Condition as predictor variables and Participant as random effect (see Figure 65). Data was normally distributed as indicated by a Shapiro-Wilk test (p =0.6198, W =0.97965). Results revealed a small (f^2 = 0.09) but significant effect of Group (p = 0.00779, SE = 0.3957) with Control group decreasing by -1.1245 units, showing that the Control group displayed significantly lower voltages than the Experimental group. The effect of Condition was not significant (p = 0.55857, SE = 0.3819), despite a slightly negative coefficient for switch condition (-0.2259). The VIF for Group and Condition was equal to 1, revealing no sign of multicollinearity. There was no significant interaction between Group and Condition (p = 0.353; SE = 0.7683)

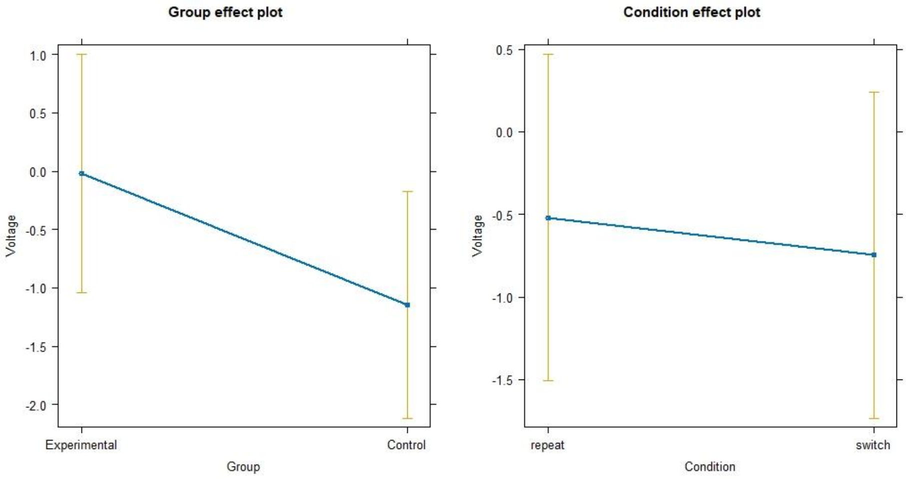


Figure 65: Linear mixed model for between-group analysis in N200 time window

As depicted in Table 21, the Experimental group showed lower switching costs at Baseline compared to the Control group. Linear mixed models that analysed the effect of Group on Voltage separately for switch and repeat conditions revealed that there was no effect of Group on Voltage in the repeat condition (p = 0.238, SE =0.5271) despite a negative correlation for the Experimental group (-0.613). Furthermore, there was only a moderately significant effect of Group on Voltage in the switch condition (p =0.0832, SE =0.75640).

Table 21: Mean voltage in switch and repeat condition per group N200 time-window

	Experimental group	Control group
Mean voltage switch N200	-0.083807, SD =2.5885	-1.420941, SD =1.849674

Mean voltage repeat N200	-0.253517, SD =1.862124	-0.865414, SD =1.242109
Switching costs N200	-0.16971	0.5555275

The finding that the Control group displayed significantly lower voltages than the Experimental group corresponds to the findings for the P300 component, where more negative voltage was also found, indicating more resource allocation to attention switching, stimulus categorization and working memory updating (see Daffner et al. (2010) and Polich (2007), Cespón & Carreiras (2020) Morrison & Taler (2020)), actually speaking for a pre-existing advantage of the SI group. However, as already mentioned, the lower voltages in the N200 and P300 time-windows coincided with lower accuracy in the Experimental group.

For the Simultaneous interpreting group, event-related potentials for all sessions for the N200 time-window show amplitudes that can be interpreted to correspond to the N200 component as depicted in Figure 66:

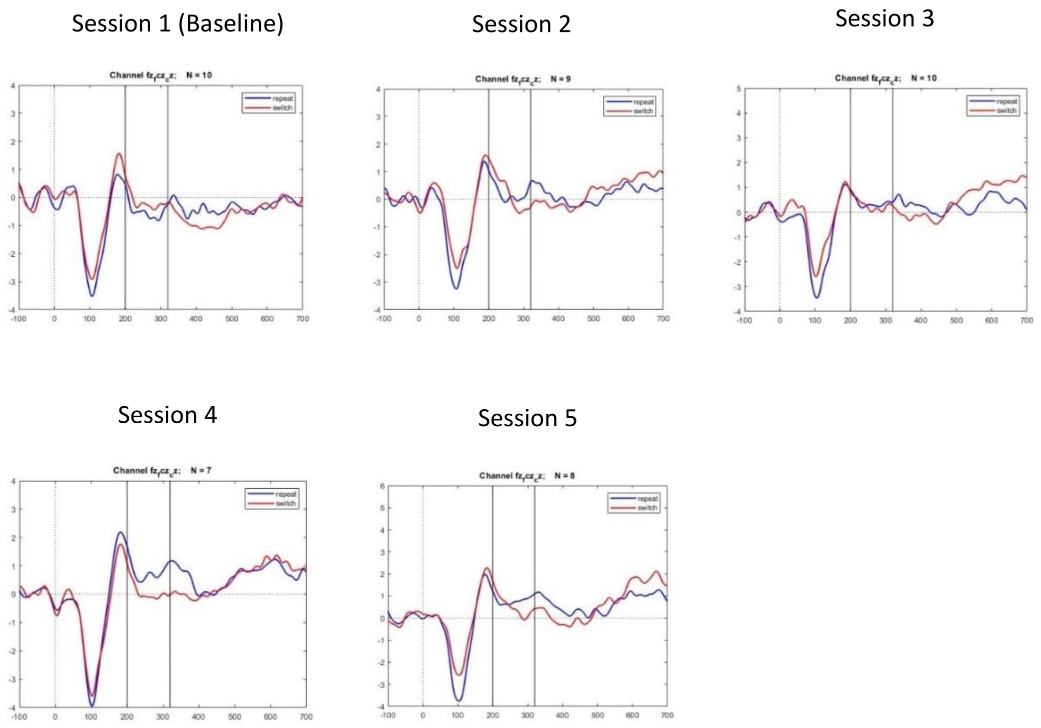

Session 1 (Baseline)
Session 2
Session 3
Session 4
Session 5

Figure 66: Event-related potentials N200 component all sessions

By means of a 2D topographical map, Figure 67 shows the distribution of the averaged voltage at the time window 200-320ms after stimulus onset on the scalp per session and per condition. Similar to the Flanker task, voltage becomes more negative (indicated in blue) in the frontal regions until Semester 3 (Session 4).

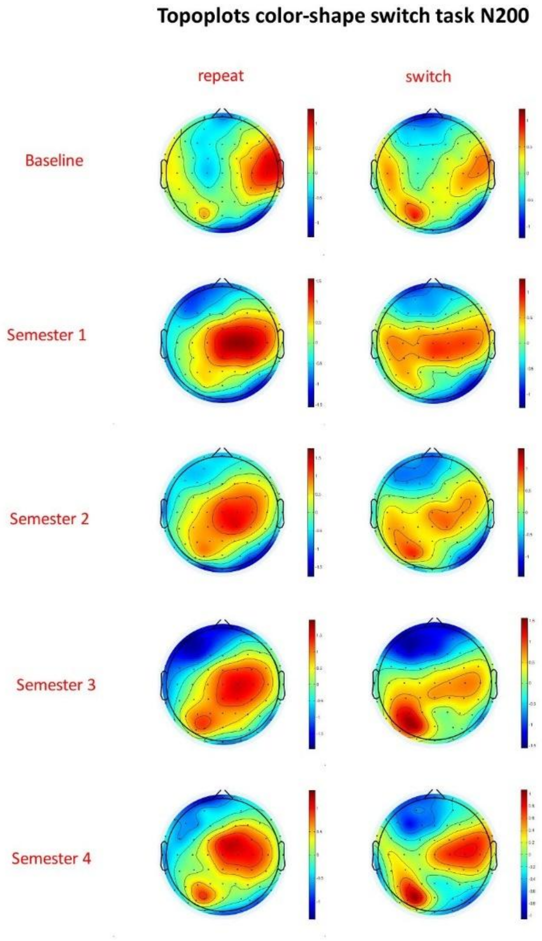


Figure 67: Topoplots colour-shape switch task in N200 time-window

Mean voltage in the N200 time-window across sessions depicted in Figure 68 reveal an initial increase in mean voltage until Session 3, followed by a drastic decrease especially for the switch condition in Session 4. For Session 5, voltage in the switch condition increases again. It seems that the repeat condition remains quite constant after initially increasing, while the switch condition is affected more by the course of the interpreting training.

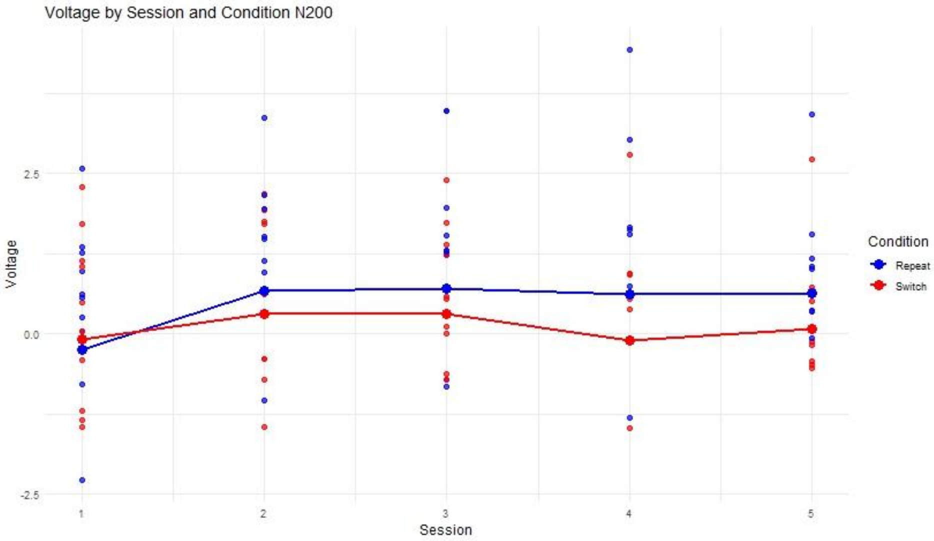


Figure 68: Progression of mean voltage in N200 time-window across sessions

Before fitting a linear model, data was confirmed to be normally distributed by running a Shapiro-Wilk test (p = 0.2495, W = 0.98169). The linear mixed model that was fitted with Voltage as dependent variable, Session and Condition as predictor variables as well as Participant as random effect (see Figure 69) revealed neither a significant effect of Session on Voltage (S2: p = 0.323, SE =0.570411, S3: p= 0.226, SE =0.552547, S4: p = 0.690, SE = 0.619766, S5: p =0.545, SE 0.593187=), nor a significant effect of Condition on Voltage (p =0.354, SE =0.372527), although the coefficient of the switch condition showed a slight decrease by -0.347381 units. VIF-values for Session and Condition were equal to 1, not indicating any collinearity. Including an interaction term to see whether there is an interaction between Condition and Session did not reveal significant results.

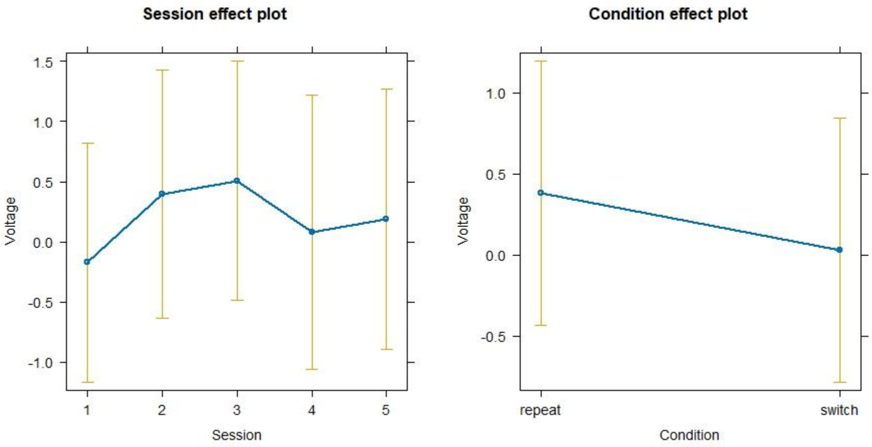


Figure 69: Linear mixed model of voltage in N200 time-window

The finding that no effect of Session on voltage was significant across sessions does contradict the hypothesis formed based on literature that has shown an improvement in behavioural performance in task switching (Babcock et al. (2017)) and Dong & Liu (2016)) but corresponds to previous findings by van de Putte et al. (2018) who did not find an improvement in inhibitory executive control after SI training, the N200 component in task switching representing the interference suppression aspect of task switching (Jamadar et al. (2015), Cespon & Carreiras (2020)).

The electrophysiological switching costs of the N200 time-window correspond to the switching costs in the P300 time-window (see Figure 70), showing an increase in switching costs until Session 4, after which they decreased again. A linear mixed model with switch costs per participant as dependent variable, Session as predictor and Participant as random effect showed a significant effect of Session 4 and Session 5 on switching costs (S4: p = 0.032, SE = 0.36, S5: p = 0.043, SE = 0.35). This mirrors the progression of voltage in both conditions in Figure 68 where the switch condition becomes more negative again especially in Session 4. The finding for significant switch costs in Session 4 and 5 that become more positive does not correspond to previous literature that has shown gains in switch costs after SI training (see Dong & Liu (2016) and Babcock et al. (2017)).

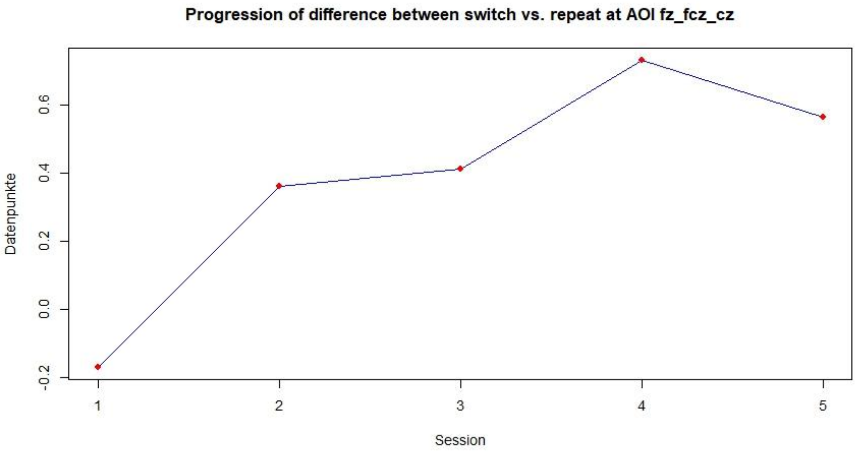


Figure 70: Electrophysiological switching costs in N200 time-window across sessions

Mirroring the peak analysis for the Flanker task, peak latencies were computed for the N200 time-window in the colour-shape switch task by computing the most negative peak in the N200 time-window for both conditions across sessions. Interestingly, Figure 71 reveals a dissociation of latencies in both conditions. While latency in both conditions increases initially in Session 2, the switch condition decreases steadily until Session 4, after which latency increases again. In contrast to that, repeat condition latency steadily increases until Session 4, which is followed by a decrease in latency. A linear mixed model did not reveal any significant effect of Session on latency (all Session $p > 0.05$) but a small ($f^2 = 0.04$) but significant effect of Condition on latency ($p = 0.047$, $SE = 5.972$) with the switch condition regression coefficient increasing by 12.074 units, revealing significantly longer latencies for the switch condition. No indication for collinearity was indicated by the VIF-values for the predictors Session and Condition (VIF = 1). Adding an interaction term to check whether the effect of condition on latency is modulated by Session did not show significant results (all $p > 0.05$).

Lastly, the finding that Session did not affect N200 latency significantly does not correspond to the expectations formed based on N200 literature that assigns later latencies with increased cognitive resource allocation (Jamadar et al. 2015), however, the finding of a significant effect of Condition on latency with longer latencies for switch condition than for the repeat condition corresponds to relevant N200 literature (Cespón & Carreiras 2020).

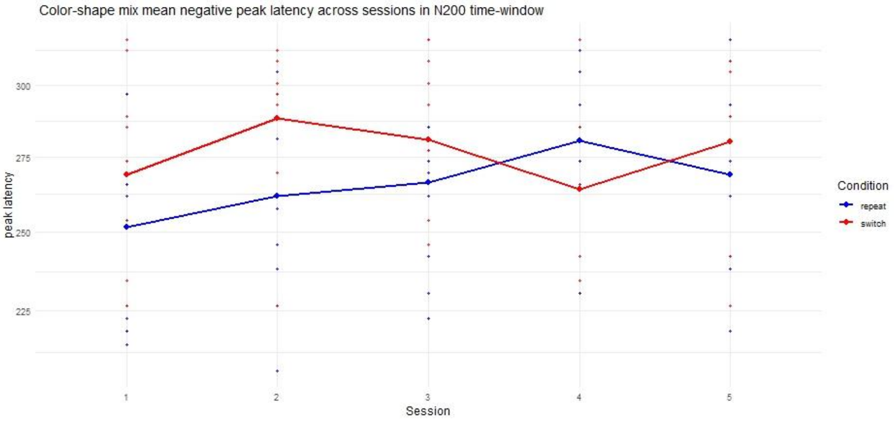


Figure 71: Peak latencies for colour-shape switch task in N200 time-window

To again depict the effect that participant variability might have had on the results of the longitudinal analysis, Figure 72 presents the progressions of mean voltage per participant and condition. Although there seems to be a trend for a positive peak at Session 3 as can be seen in Figure 68 and Figure 69, the variability between participants is very high.

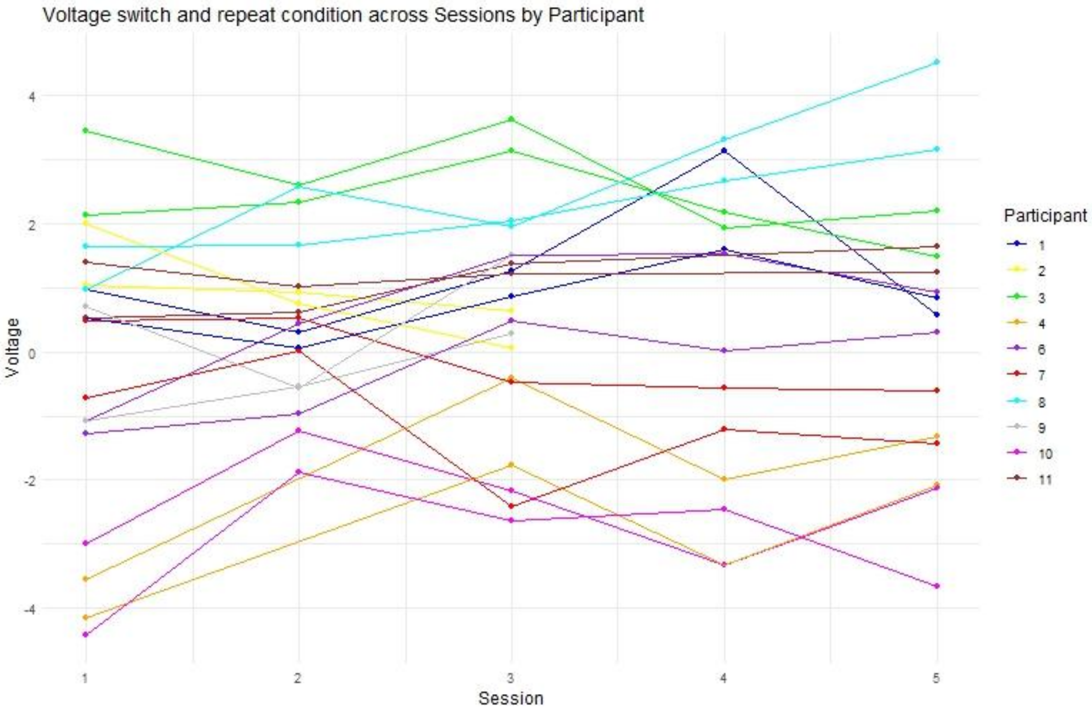


Figure 72: Progression of mean voltage in the N200 time-window across sessions per participant and condition

8.2.3.3 2-back

Voltages of the Experimental and the Control group were compared at baseline level in the P300 time-window by fitting a linear mixed model with Voltage as outcome variable, Group and Condition as predictors and Participant as random effect. Data was normally distributed as indicated by a Shapiro-Wilk test (p = 0.2036; W = 0.96392). Results did not show a significant effect of Group on Voltage (p =0.163, SE = 0.4481) mirroring the reaction time results. However, the effect of Condition was large ($f^2 = 0.37$) and highly significant (p < 0.001, Se = 0.2477) with a decrease by -1.5117 units of the coefficient for the no-match condition (see Figure 73). VIF-values were equal to 1 for both Session and Condition. There was no significant interaction between Group and Condition (p = 0.6334; SE = 0.5103).

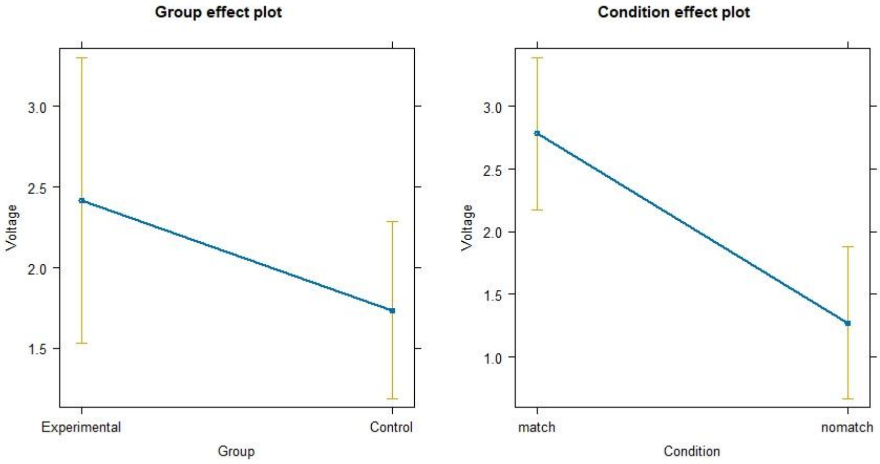


Figure 73: Linear mixed model of between group comparison at baseline 2-back task

Table 22 depicts the voltage per group and per condition.

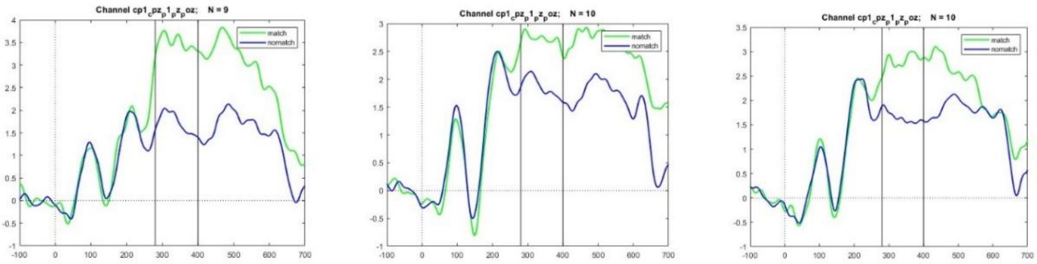
Table 22: Descriptive statistics 2-back Experimental and Control group

	Experimental group	Control group
Mean voltage match	3.173, SD = 1.905	2.439, SD = 1.356
Mean voltage no-match	1.519, SD = 0.984	1.033, SD = 0.569

Separate linear mixed models were fitted per condition to see if there is a difference between groups. However, neither the effect of Group in the match condition (p = , 0.246 SE = 0.6340), nor the effect of Group in the no-match condition was significant (p = 0.143, SE =0.3192).

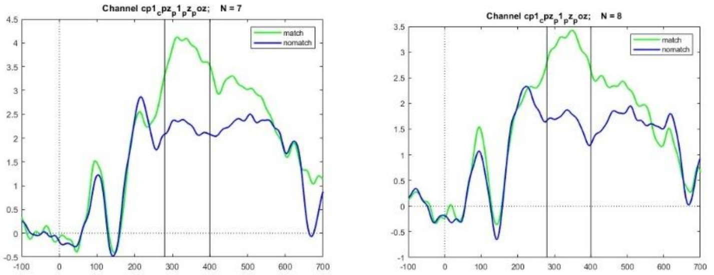
The finding that there is no significant difference between groups at baseline in the P300 time-window corresponds to the present reaction time results but not to the present accuracy results that showed higher accuracy for the Experimental than for the Control group. However, the present ERP results correspond to previous literature of behavioural studies that did not find a difference in WM-updating performance between groups at baseline (see e.g. Dong & Liu (2016), Rosiers et al. (2019) and Babcock et al. (2017). The result that Condition had a significant effect on Voltage with the no-match condition showing lower voltage than the match condition corresponds to results in relevant n-back literature (see Barker & Bialystok (2020)).

Figure 74 depicts event-related potentials for the longitudinal study with the interpreting group per session. Within the P300 time-window, an amplitude that fits the characteristics of a P300 amplitude can be detected.


Session 1 (Baseline)

Session 2

Session 3


Session 4

Session 5

Figure 74: Event-related potentials for the 2-back task across sessions

By means of a 2D topographical map, Figure 75 shows the distribution of the averaged voltage at the time-window 280-400 ms after stimulus onset on the scalp per session and per condition. Especially interesting is the increase in positive voltage (indicated in red) in the parietal region for the match condition until Semester 2 (Session 3), after which voltage seems to decrease again in this region.

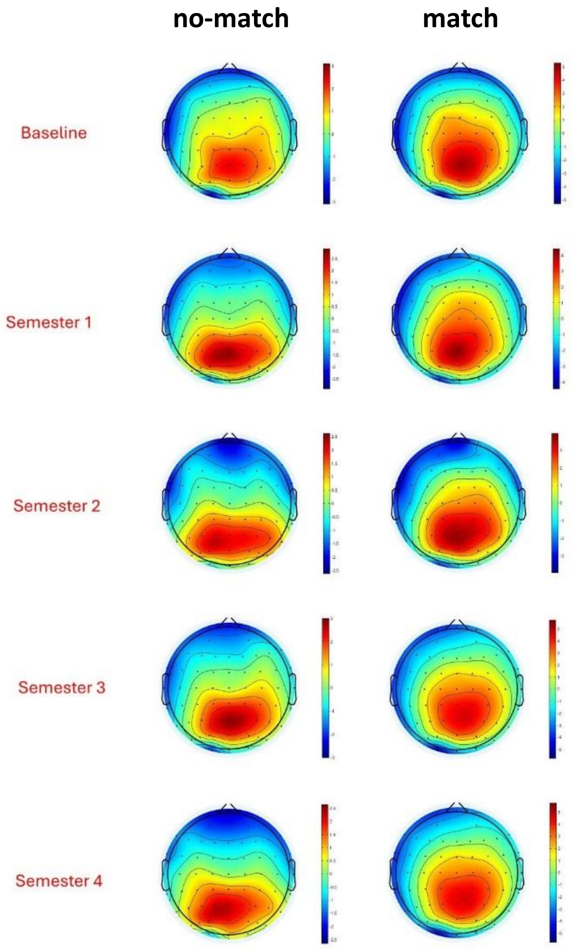


Figure 75:Topoplots for 2-back task

Figure 76 depicts the progression of mean voltage for the Experimental group across sessions and per condition. Mean voltage in the match condition is always more positive than in the no-match condition, which corresponds to literature postulating larger P300 amplitudes for matching stimuli to previous stimuli that are recognized as different or important and therefore updating memory representations (see

Barker & Bialystok (2019)). Interestingly, this ability seems to decrease from baseline to Session 2, which is then followed by a steady increase until a positive peak at Session 4, followed by a slight decrease in mean voltage for Session 5. This becomes apparent in both conditions, although more pronounced in the match condition.

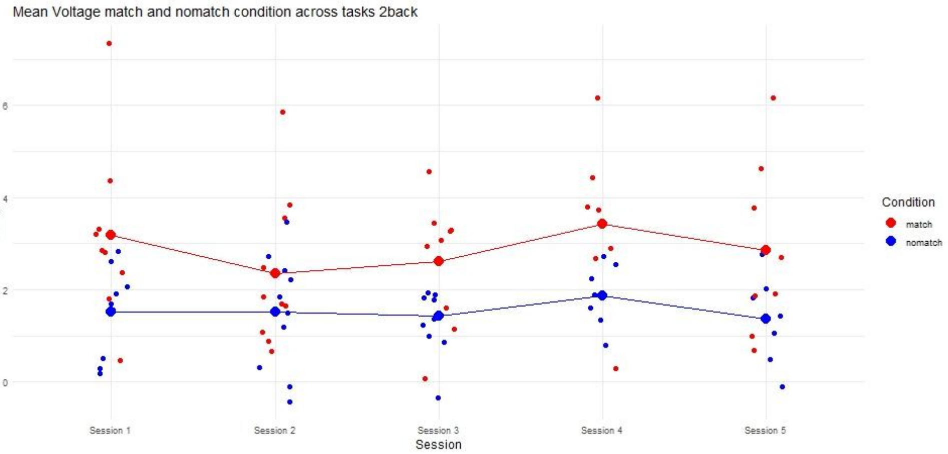


Figure 76: Progression of mean voltage across sessions per condition for the 2-back task

A linear mixed model with Voltage as outcome variable, Session and Condition as predictor variable as well as Participant as random effect was fitted to check for the effect of Session and Condition on Voltage (see Figure 77). Data showed normal distribution as indicated by a Shapiro-Wilk test (p = 0.227; W = 0.98108). None of the effects of Session on Voltage were significant (S2: p =0.423, SE =0.4094; S3: p =0.539, SE = 0.4094; S4: p= 0.597, SE =0.4510; S5: p = 0.563, SE = 0.4327), although the predictor coefficients indicated a decrease by -0.3297 units for Session 2, a decrease by -0.2524 units for Session 3, an increase by 0.2394 units for Session 5 and another decrease by -0.2511 units for Session 5. The effect of Condition on Voltage was medium-sized (f^2 = 0.23) and highly significant (p < 0.001, SE =0.2670) with a decrease of the predictor coefficient of the match condition by -1.3144 units, revealing significantly lower voltage for the no-match condition. There was also no collinearity between the predictors as indicated by VIF-values equal to 1 for Session and Condition. There was no significant interaction between Session and Condition (all p> 0.05). There was no statistically significant interaction between Session and Condition.

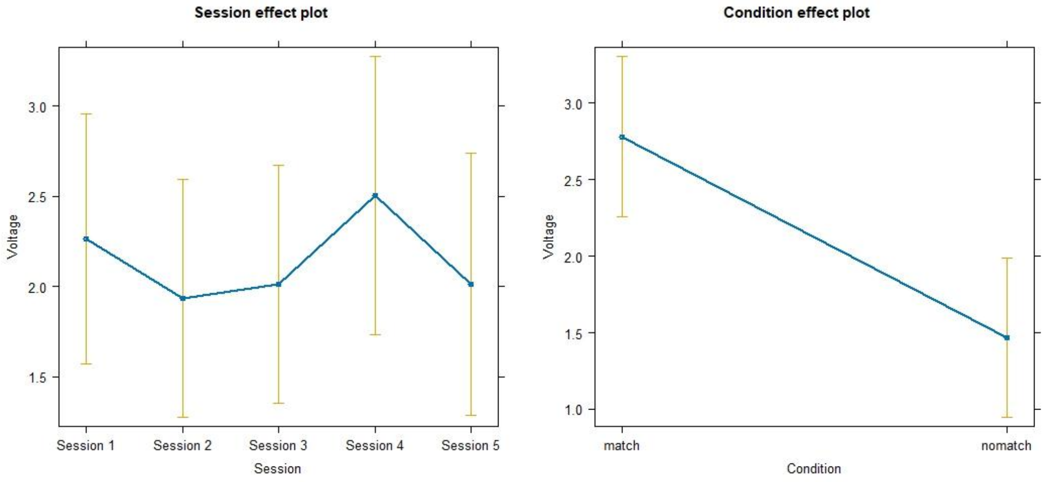


Figure 77: Linear mixed model for the Experimental group across sessions

The finding that none of the Sessions had a significant effect on Voltage in the P300 time-window across all Sessions contradicted the present behavioural literature as well as previous literature that has found a significant effect of post-test on reaction time or accuracy in working memory measures (see Dong & Liu (2016) and Babcock et al. (2017)).

For the 2-back latency analysis, the most positive peak within the P300 time-window was computed for both conditions across sessions. Figure 78 reveals a bridge-shaped progression of latency in the no-match condition, showing that latency increases initially and then decreases after Session 3. Latency in the match condition also increased initially and then drastically decreases for Session 3, followed by another decrease from Session 4 to Session 5. A log transformed linear mixed model with peak latency as dependent variable and Session and Condition as predictor variables as well as Participant and random effect revealed only a significant effect of Session 2 on latency (p = 0.024, SE = 0.029). Session 3-5 did not show significant effect of Session on latency (S3: p = 0.106, SE = 0.047, S4: p = 0.216, SE = 0.039; S5: p = 0.850, SE = 0.030). The overall main effect of Session was small- to medium-sized (f^2 = 0.09). The effect of Condition on latency was not significant (p = 0.812, SE = 0.019) which does not correspond to existing n-back literature such as by Barker & Bialystok (2019), who found longer P300 latencies on match than on no-match trials for the 1-back task. VIF-values did not indicate any collinearity between predictors (VIF = 1).

Furthermore, no significant interaction effect between Session and Condition could be observed.

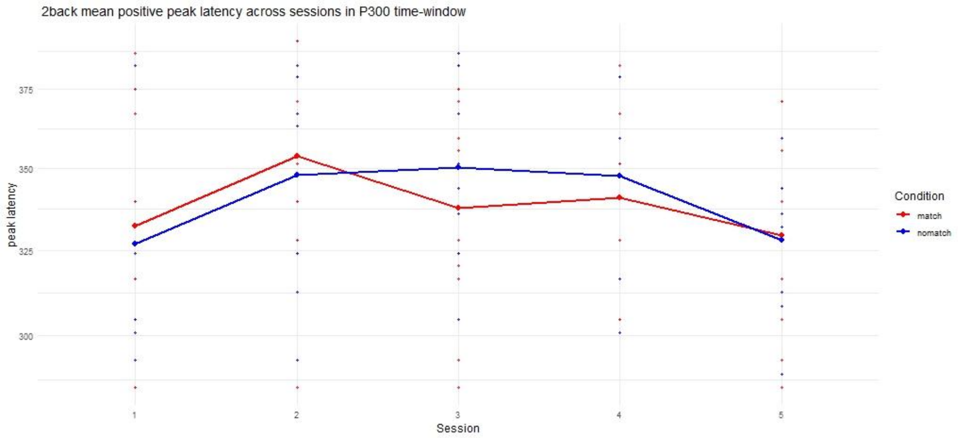


Figure 78: Peak latency of 2-back task in P300 time-window

8.2.3.4 3-back

To compare the Experimental and the Control group at Baseline on voltage within the P300 time-window, a linear mixed model with Voltage as outcome variable, Group and Condition as predictor variables and Participant as random effect was fitted (see Figure 79). Data showed normal distribution as indicated by a Shapiro-Wilk test (p = 0.9279, W =0.98821). Results showed no significant effect of Group on voltage (p = 0.11285, SE = 0.3053) but a medium- to large sized (f^2 = 0.2) and significant effect of Condition on Voltage (p = 0.00168, SE =0.2981) with the no-match coefficient decreasing by -1.0274 units from the match condition, indicating significantly lower voltages for the no-match condition. VIF-values for Session and Group were equal to 1 and indicated no collinearity between predictors. No significant interaction between Group and Condition could be observed (p = 0.37, SE = 0.60).

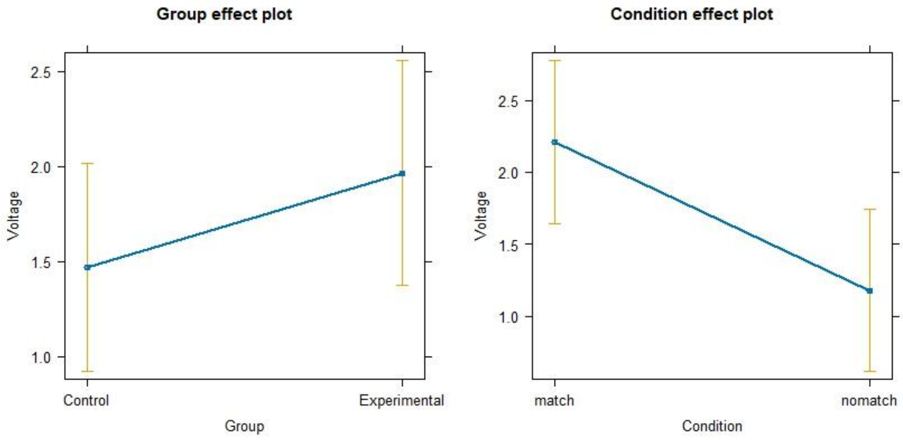


Figure 79: Linear mixed model for between-group comparison of Voltage for P300 time-window

Table 23 shows mean voltages for both groups and conditions at baseline in the P300 time-window:

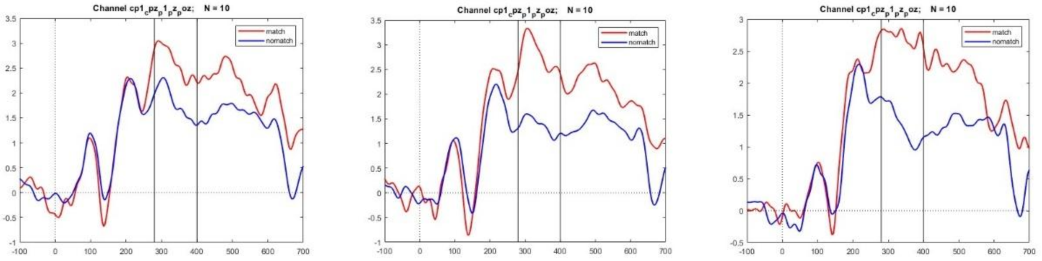
Table 23: Mean voltage per condition and per group 3-back task

	Experimental	Control
Mean voltage match	2.271, SD = 1.654	2.11, SE = 1.37
Mean voltage no-match	1.541, SD = 0.694	0.831, SE = 0.725

The finding that there is no significant difference between groups at baseline in the P300 time-window corresponds to the present behavioural results as well as to previous literature of behavioural studies that did not find a difference in WM-updating performance between groups at baseline (see e.g. Dong & Liu (2016), Rosiers et al. (2019) and Babcock et al. (2017). Furthermore, the result that Condition had a significant effect on Voltage with the no-match condition showing lower voltage than the match condition corresponds to results in relevant n-back literature (see Barker & Bialystok (2020)).

Figure 80 shows the event-related potentials for the experimental group per sessions. Within the time-window 280-400ms, an amplitude resembling the P300 component can be seen.

Session 1 (Baseline) Session 2 Session 3



Session 4 Session 5

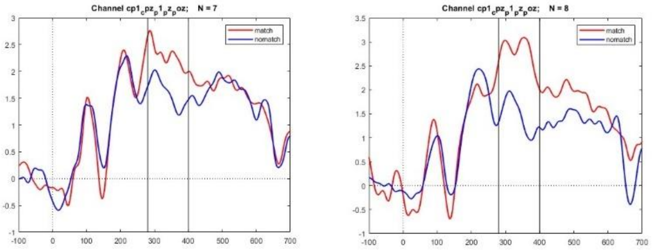


Figure 80: Event-related potentials for P300 time-window in 3-back task across sessions

The distribution of the averaged voltage at the time-window 280-400 ms after stimulus onset on the scalp per session and per condition can be seen in Figure 81 in a 2D topographical map. As for the 2-back task, an increase of positive voltage in the parietal area becomes apparent for the match condition until Session 3, after which voltage in this area seems to decrease slightly. Strikingly, voltage in the no-match condition becomes more positive and spreads more to lateral areas consistently after baseline until Session 5.

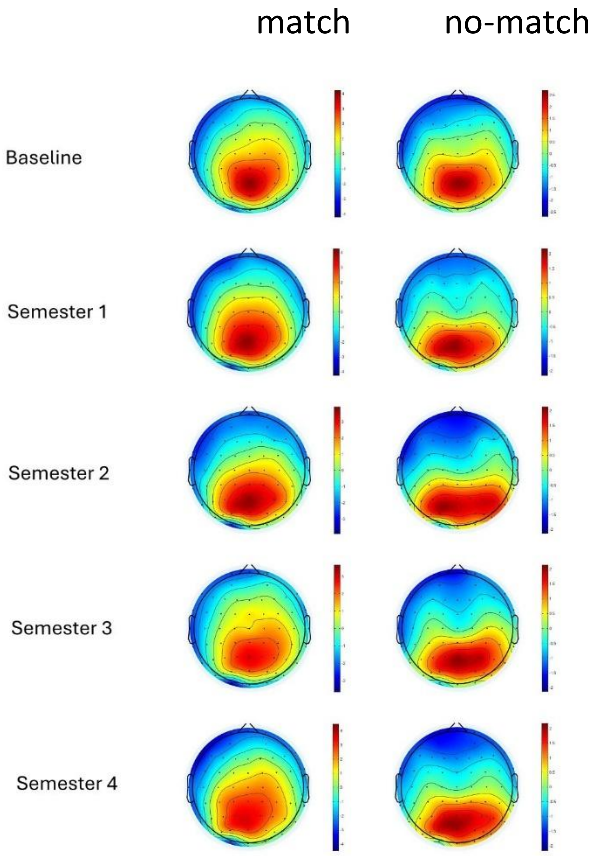


Figure 81: Topoplots for 3-back task in both conditions across sessions

The progression of mean voltage per session and condition (see Figure 82) shows a diverging pattern. While voltage in the match condition increases until Semester 3, after which is reaches a negative peak in Session 4 only to increase again in Session 5, voltage in the no-match condition initially decreases after Session 1, which is followed by a steady increase until Session 4 and a small decrease for Session 5.

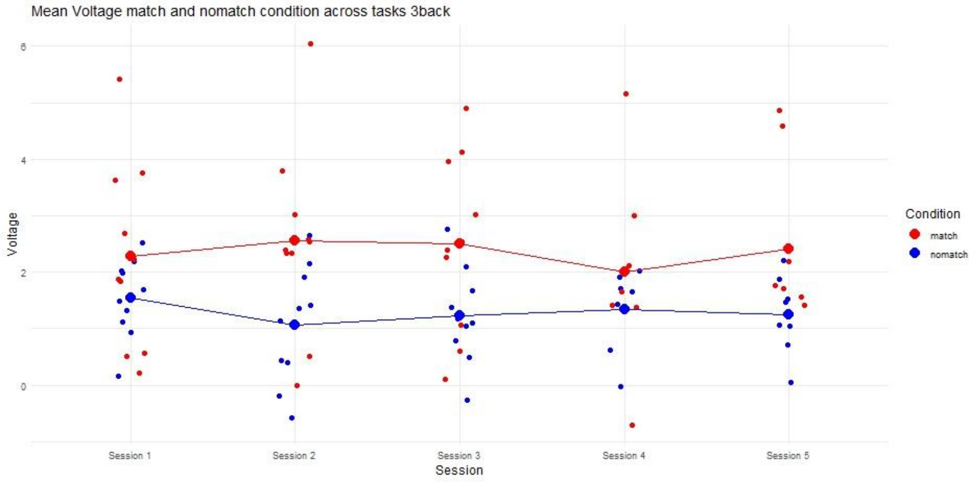


Figure 82: Progression of mean voltage in P300 time-window across sessions for 3-back task

Subsequently, a linear mixed model with Voltage as outcome variable and Session and Condition as predictor variables as well as Participant as random effect was fitted (see Figure 83). A Shapiro-Wilk normality test revealed that residuals were not normally distributed (p = 0.01666, W = 0.9653). Since log transformation could not be applied because of the negative values from the EEG data, a Yeo Johnson transformation was performed. In opposition to log transformation, the Yeo Johnson transformation allows for the transformation of nonpositive data. However, the Yeo Johnson transformation also did not solve the issue of non-normally distributed data and did not alter the results notably. Results did not show any significant effect of Session on Voltage, which might be due to the big variance of datapoints as visible in Figure 82 (S2: p =0.781 , SE = 0.34611; S3: p = 0.892, SE = 0.34611; S4: p = 0.373, SE = 0.38869; S5: p = 0.712, SE = 0.37193). Despite that, the effect of Condition on Voltage was medium-sized (f^2 = 0.18) and highly significant (p < 0.001, SE = 0.23074) with the coefficient of the no-match condition decreasing by -1.08722 units compared to the match condition, showing significantly higher voltages for the match than the no-match condition (see Figure 83). No collinearity was indicated by the VIF-values for the predictors Session and Condition (VIF = 1). After including an interaction term to see whether the effect of Condition on Voltage is modulated by Session, no significant effects were revealed (all p < 0.05).

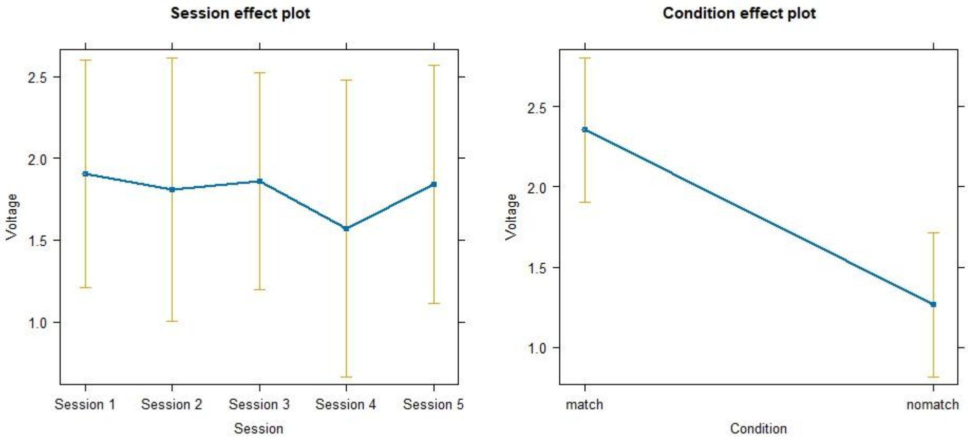


Figure 83: Linear mixed model for P300 time-window in 3-back task across sessions

The finding that there was no effect of Session on Voltage in the P300 time-window across Sessions did not correspond to the present behavioural literature that showed faster RT and increased accuracy over testing. Furthermore, it contradicts previous literature that has found a significant effect of post-test on reaction time or accuracy in working memory measures (see Dong & Liu (2016) and Babcock et al. (2017)). The significant effect of Condition with lower voltage for the no-match condition however is in line with previous n-back literature (see Barker & Bialystok (2019)).

As in the 2-back task, the peak latency in the P300 time-window was computed for both conditions of the 3-back task across sessions. Figure 84 reveals that while latencies of the match and no-match conditions both initially decrease, latency of the match condition increases after Session 2 while latency of the no-match condition further decreases. Latency of conditions then meet at Session 4 at almost the same level. This variability is also reflected in the linear mixed model: None of the effects of Session on latency are significant (all Session $p > 0.05$), neither is the effect of Condition on latency ($p = 0.920$, $SE = 3.127$). This is not in line with previous n-back literature that associates longer latencies with the match condition (see Barker & Bialystok (2019)). As indicated by residuals, data was not normally distributed. A log-transformed linear mixed model was fitted, that however did not reveal any substantially different effects and did not remedy the deviation from normality (Session: all $p > 0.05$); Condition: $p > 0.05$). VIF-values did not indicate any collinearity between predictors (VIF = 1). Finally, no interaction between Session and Condition was observed.

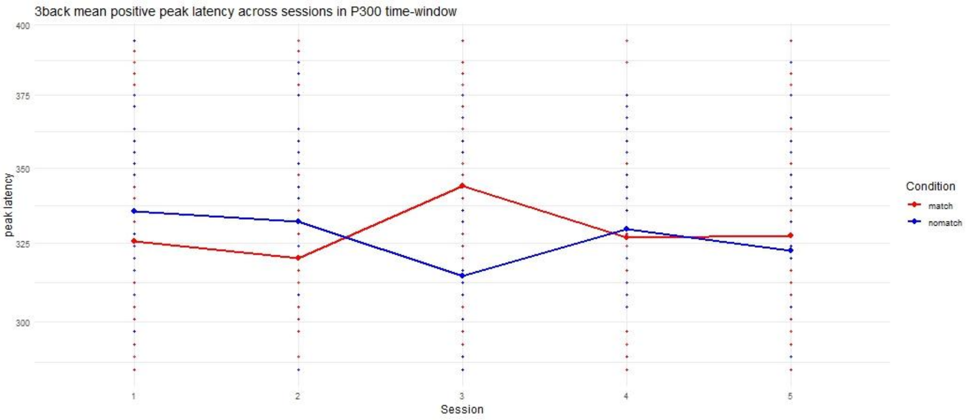


Figure 84: Peak latency 3-back task in P300 time-window

8.2.3.5 2-back vs. 3-back task

Finally, how the task 2-back and 3-back task as well as Session affected the voltage in the P300 time-window across participants was analysed by fitting a linear mixed model with voltage as dependent variable, Task (2-back/3-back), Condition (match/no-match) and Session as independent variables and Participant as random effect was fitted. Data did not show normal distribution which is why the data was initially transformed to approach a normal distribution. Since log transformation could not be applied because of the negative values from the EEG data, a Yeo Johnson transformation was performed, as the Yeo Johnson transformation allows for the transformation of nonpositive data. However, after performing the Yeo Jonson transformation, data normality could not be improved since data still did not show normal distribution as indicated by a Shapiro-Wilk test (p = 0.01377, W = 0.98056). However, robust linear mixed models can handle a certain degree of non-normality.

Results revealed a medium- to large-sized (f^2 = 0.2) significant effect of Condition on Voltage (p > 0.001, SE = 0.11808) with the no-match condition showing more negative voltage in the P300 time-window than the match condition, indicated by the coefficient of the no-match condition that decreased by -0.78637 units. Contrary to the expectation, the effect of Task on Voltage was small ($f^2 = 0.02$) and only marginally significant (p = 0.0645, SE = 0.11823), although there is a tendency for the 3-back task to show lower voltage than the 2-back task, indicated by the 3-back task coefficient that decreased by -0.22008 units. There was no significant main effect of Session on Voltage (all p > 0.05). Results are depicted in Figure 85. VIF-values for the predictors Condition, Task and Session were all equal to 1, not indicating any multicollinearity.

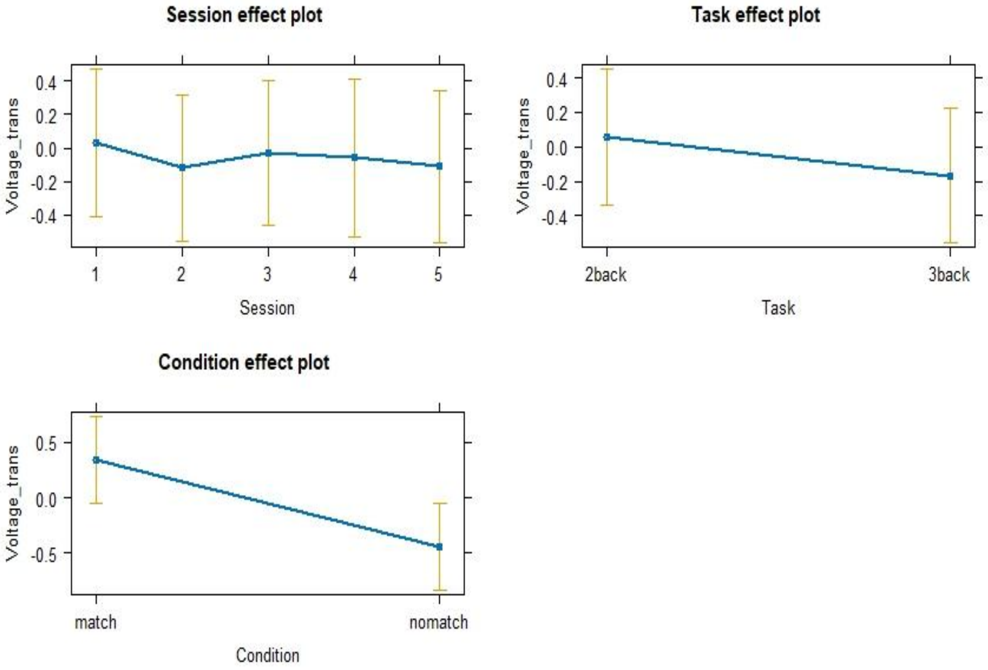


Figure 85:Linear mixed model for comparison of voltage in 2-back and 3-back task across sessions

A 2-way interaction term was added to see whether Task is modulated by Session and whether Condition is modulated by Session and whether Condition is modulated by Task. Furthermore, a three-way interaction term was added to see whether Condition is modulated by Task and Session. Results did not reveal any significant interactions (all p > 0.05).

The significant effect of Condition on Voltage with the no-match condition showing more negative voltage in the P300 time-window than the match condition corresponds to previous n-back literature (see Barker & Bialystok 2020). Furthermore, although only marginally significant, the effect of Task on voltage with the 3-back task showing lower voltage than the 2-back task is in line with the present behavioural results that show longer reaction times and lower accuracy for 3-back than for 2-back task. The effect is furthermore in line with relevant ERP-literature that associates increasing memory load and task difficulty with decreasing P300 voltage and amplitudes (Kok et al. (2001), Cespon & Carreiras (2020), Daffner et al. (2011), Morrison & Taler (2020), Dong et al. (2015)).

8.2.4 Frequency analysis

Based on the theoretical framework presented in Section 8.1.2, time-frequency analysis was carried out for all three tasks. This was done in the EEGlab environment where it is possible to examine event-related spectral perturbations (see Section 8.1.2). However, the EEG lab environment only allows for single-participant analysis. Therefore, and since the frequency analysis was not the main focus of this dissertation, only single participants were examined by visual inspection, which makes the results anecdotal and only allows for a hypothetical interpretation of results. No correlation can be drawn for the whole sample. Nevertheless, since all sessions were compared in one participant, the results can give some interesting insights from the frequency domain into how executive functions progress over time.

Before getting into the analysis, it has to be mentioned that ERSP usually describe the perturbations from baseline. Since no data from a “task-absent” condition was gathered, the only option is to compare the congruent and incongruent conditions which is not entirely sound. Furthermore, there is the possibility that pre-stimulus cognition already affects how the alpha behaves after stimulus-presentation, which was not examined here.

As an example before discussing the analysis, Figure 86 depicts the ERSP graph for the congruent condition from the Flanker task in Session 1 (baseline) for P01. On the left y-axis, the frequency range is represented. For all tasks, the alpha frequency range (8-12 hz) was of primary interest. The x-axis shows the time in ms from -200ms before stimulus onset until 700ms after stimulus onset. The ERSP bar on the right side shows the event-related spectral perturbations in db. As explained in Section 8.1.2, high db (indicated in the red colour on the right side in the ERSP scale) represents an increase in alpha power and an increase in synchronized neuronal activity. Alpha synchronization (increased alpha power) has been associated with increased inhibitory control and less active processing (see Klimesch et al.). The following analysis will include zoomed in versions of the time-frequency plots for better visualization.

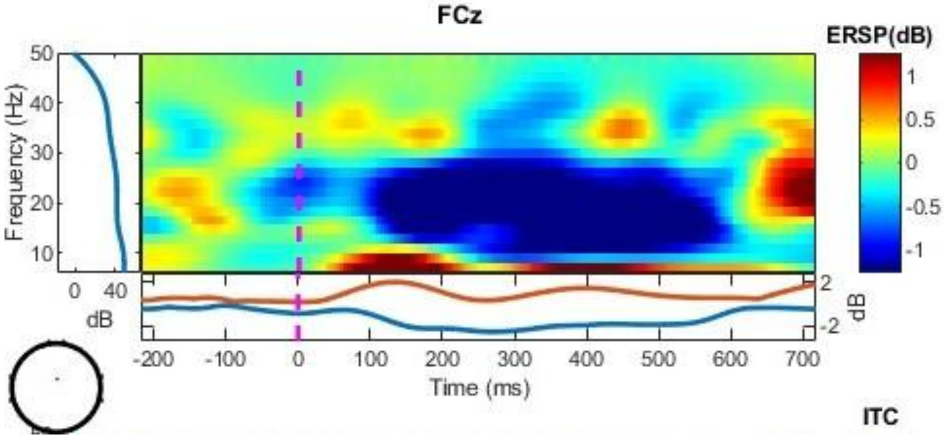
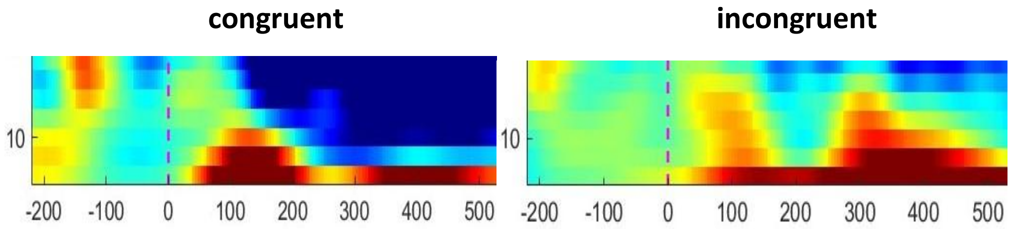


Figure 86: Example of time frequency plot

For the analysis of the Flanker task, data from P01 from both congruent and incongruent conditions were analysed through visual inspection. Alpha frequency from 8-12 hz was examined at the electrode fcz in the N200 time window from 200-400 ms. In Chapter 7, I hypothesized that over the course of testing, especially the incongruent condition will show less alpha power, so less synchronization, since less inhibitory control will be necessary over time to inhibit the distracting arrows in the Flanker task as inhibitory control has been automatized alongside language control during the intense interpreting training.

Figure 87 represents zoomed in versions of the frequency analysis plots for the congruent and the incongruent conditions across all sessions:



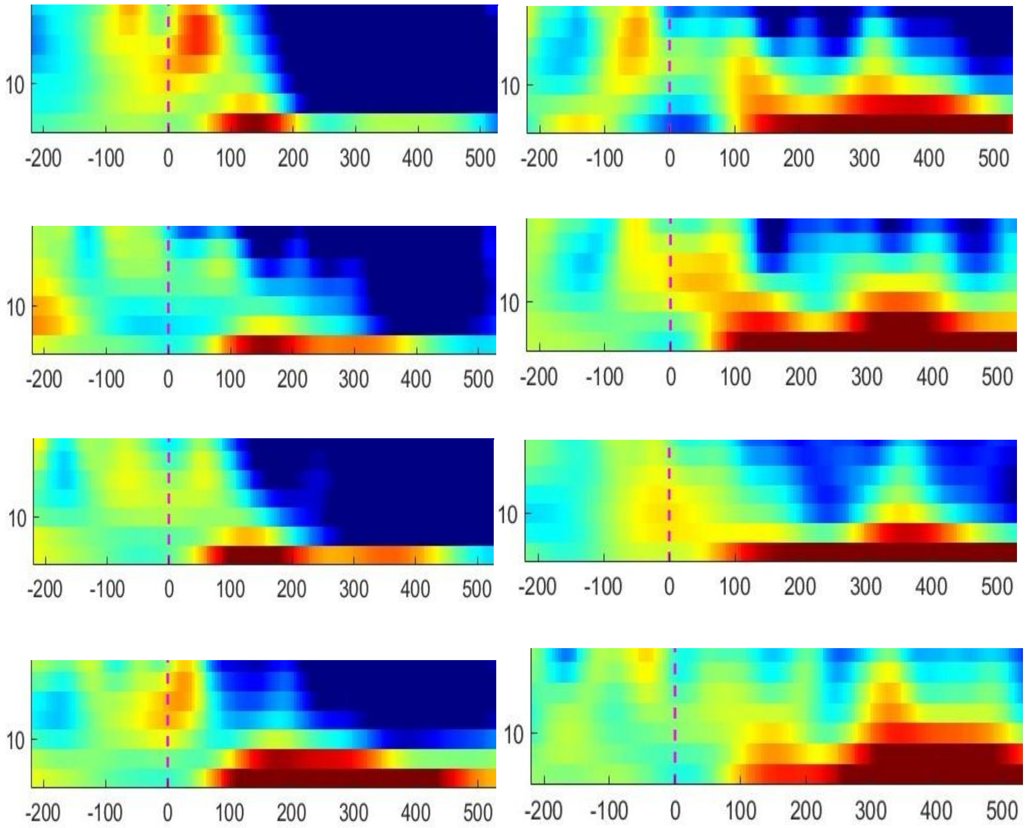


Figure 87: Time frequency plot for Flanker task across sessions and conditions

From visual inspection of the above depicted data, it becomes apparent that alpha power differs between conditions, with the incongruent condition showing more alpha power in the alpha frequency range especially around 300ms, which lies within the N200 time-window. Since according to literature, the incongruent condition involves more inhibition and conflict monitoring which should be reflected in increased alpha power, this is compliant with the theory presented in Section 8.1.2. As for the changes in alpha power across sessions, it can be observed that alpha power in the incongruent condition in the 200-400ms time-window increases until Session 3 after which it slightly decreases only to increase again at Session 5. This mirrors the progression of voltage in the ERP-analysis, where voltage decreased until Session 3. Both observations indicate increased inhibition or conflict monitoring. In the ERP-analysis, voltage increases again at Session 4 which speaks for less inhibition or conflict monitoring processes being included in solving the task which fits together with the alpha power seemingly decreasing slightly in Session 4. However, behavioural results also indicated a decrease in accuracy at Session 4. So, it is possible that less inhibitory control was engaged in Session 4 but at the cost of accuracy.

For the colour shape switch task, participant P01 was examined, and two different electrodes were chosen to target frontal activity and parietal activity, the fcz as a fronto-central electrode and the cpz as parietal electrodes. As explained in Section 8.1.2, both centro-fontal regions as well as parietal regions are relevant to the colour-shape switch tasks. The N200 component in the centro-fontal regions has been associated with inhibitory control and conflict monitoring, and the P300 component in the parietal regions has been associated with updating of task rules and allocation of attentional resources. The focus of the time-frequency plot of the fcz electrode lay in the N200 time-window from 200-320 ms after stimulus onset and the focus of the time-frequency plot of the cpz electrode lay in the P300 time-window from 280-400ms after stimulus-onset.

Figure 88 shows the time-frequency plots for P01 at the fcz electrode in the congruent and incongruent condition across all sessions. Especially in the incongruent condition is becomes apparent that alpha power increases over Session 2 and 3 after baseline (Session 1), but decreases again at Session 4 and Session 5, indicating an initial increase in required inhibitory control followed by a decrease which might indicate that in the late stage of training, less inhibitory control is necessary due to skill normalization. However, this pattern stands in stark contrast to the whole-group analysis of the N200 component in the previous section, where voltage in the N200 time-window showed a negative peak at Session 4. Since more negative N200 amplitudes have been associated with allocating more resources to inhibitory functions, this does not correspond to the alpha time-frequency pattern, for which an increase in alpha power and an increase in db (indicated by red colour) has been associated with increased inhibitory control effort. However, as depicted below, Sesson 4 and 5 show less alpha power which therefore indicates less cognitive resources allocated to inhibitory control. Since this is the data from one single participant, discrepancies from the whole sample are to be expected.

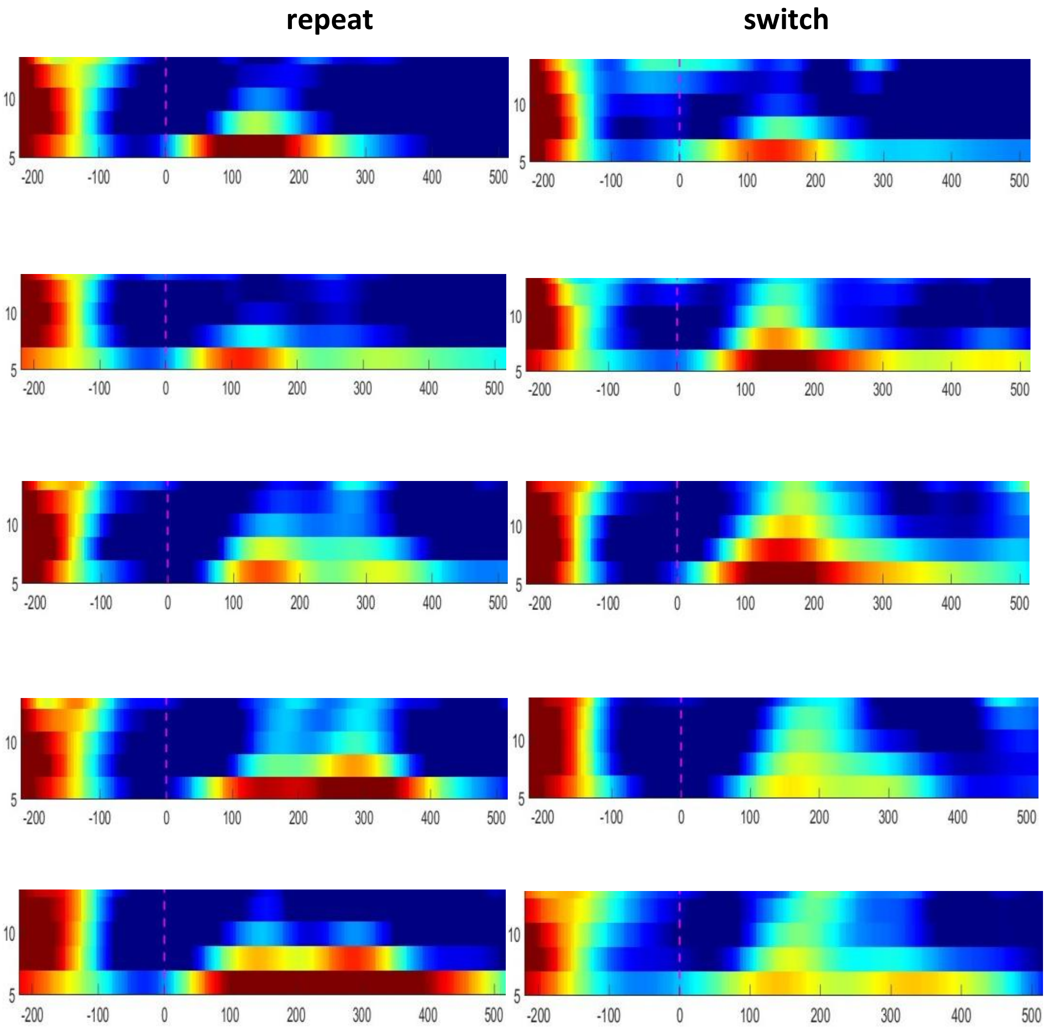


Figure 88: Time frequency plot for the colour-shape switch task at fcz across sessions and conditions

For the centro-parietally-oriented electrode cpz, Figure 89 represents the time-frequency plots for both conditions across sessions. However, no relevant alpha power changes in the P300 time-window can be observed. However, in the ERP-results for the P300 time-window, a change in voltage over Sessions could be observed and the effect of Congruency and Session on voltage was significant for Session 3 and 4. Here, one could tentatively interpret some alpha power increase in the switch condition in the lower alpha frequency range from baseline to Session 2 in the early 200-380ms time-window that decreases from Session 3 until Session 5. This trend can partly be observed in the ERP-data, where voltage increases in the P300 time window and peaks at Session 4, increased voltage in the context of the P300 component being associated with less allocation of attentional resources or less cognitive control necessary to categorize the stimulus. Here, at Session 4, alpha power is extremely low, which also speaks for less involvement of inhibitory processes. With accuracy increasing and also peaking at Session 4 this seems plausible.

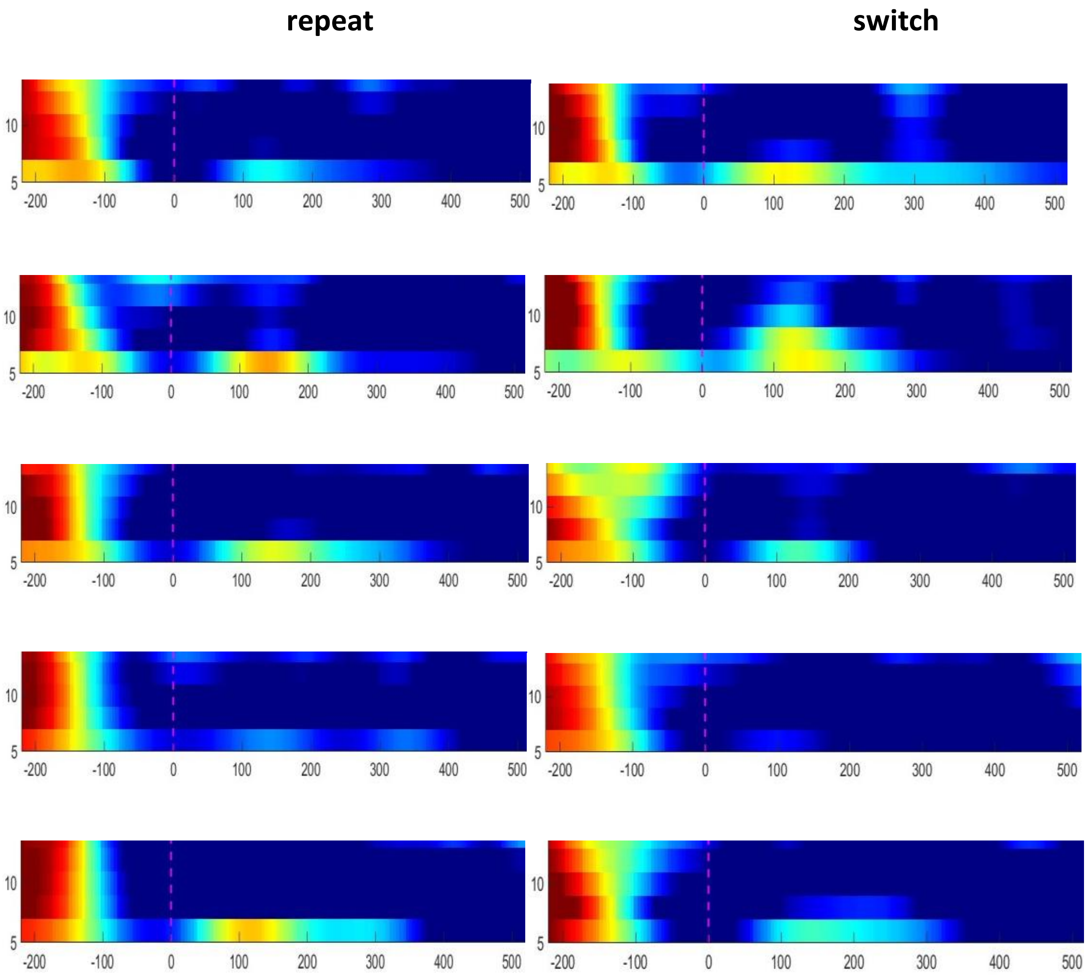


Figure 89: Time-frequency analysis for P01 at cpz for colour-shape switch task across session and condition

For the 2-back task, the centro-parietal electrode cpz was examined in P01 in the match and no-match condition across all sessions, as depicted in Figure 90.

What is striking here is that there seems to be some meaningful variability in alpha power around the 100 ms mark after stimulus onset, marked by a considerable increase in alpha power from baseline to Session 2 in the no-match condition compared to the match condition. Since increased alpha power represents increased inhibitory control effort and less active processing, this could indicate an early process of stimulus categorization. In the 2-back task, participants had to press a key when a presented letter matched a letter presented 2 trials before, otherwise, they were told to withhold a response. Speculatively, this early stimulus categorization into a “no-go” category could be the source of the increased alpha power in the no-match condition.

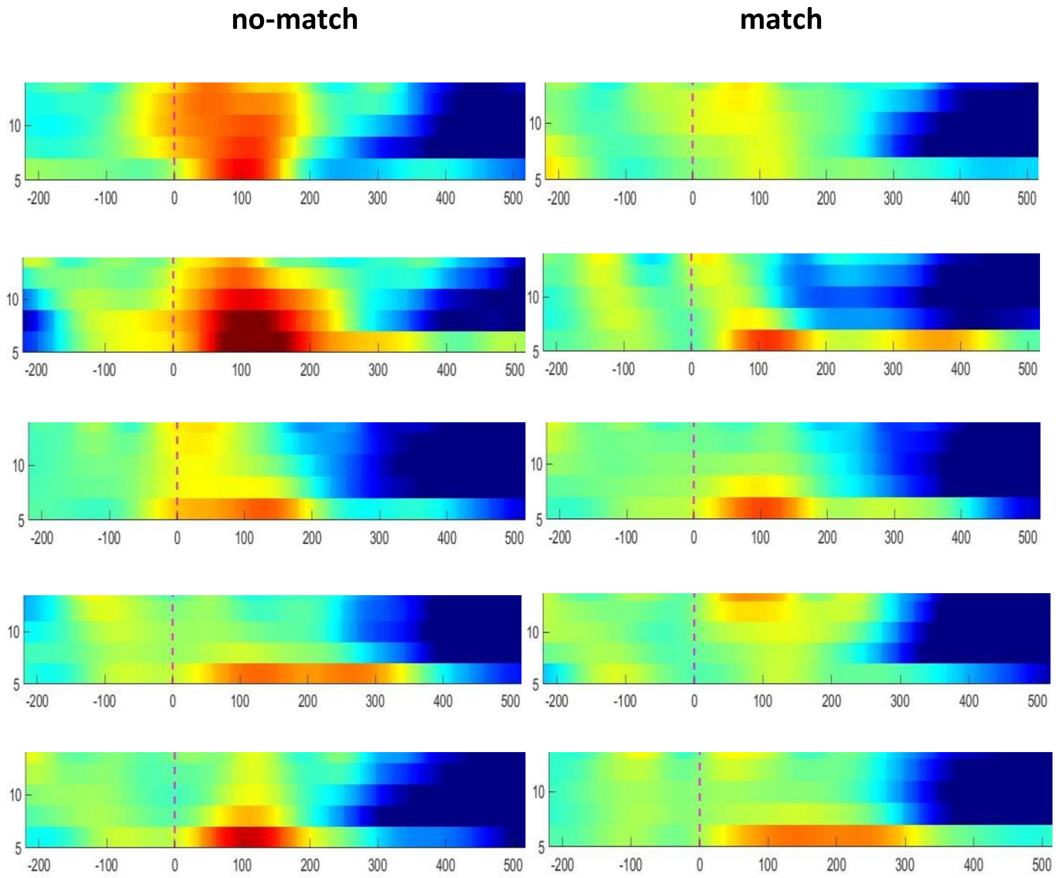
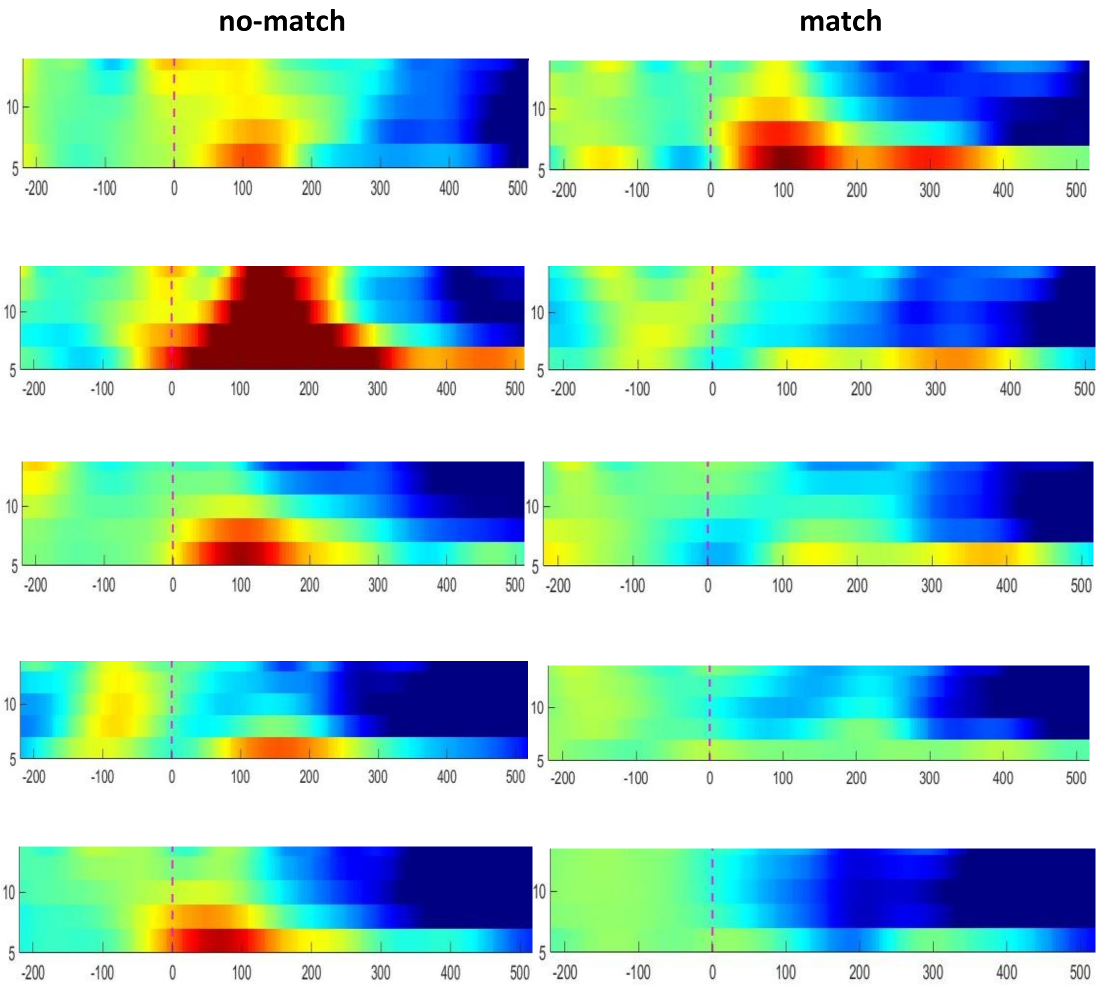


Figure 90: Frequency band plots for 2-back task across sessions and conditions

Similar to the 2-back task, the alpha power in the alpha frequency range was examined at the electrode cpz in participant P01 in the 3-back task (see Figure 91). Again, there seems to be more obvious change in the alpha power at 100ms after stimulus onset especially for the no-match condition, which might indicate early stimulus categorization activity or response monitoring to withhold the response. In the 280-400ms time point, there is some power change in the lower alpha around 8hz with alpha power decreasing especially in the match condition. Together with the increasing accuracy in the match condition from the behavioural analysis, this might indicate that less resource allocation to executive functions is necessary to solve the task of finding a match in the later sessions.


no-match

match

Figure 91: Time frequency plots for the 3-back task across sessions and conditions

9. Discussion

All results are now discussed against the backdrop of the research question of how different executive functions change as a result of interpreting training and also in the context of the relevant literature discussed in the previous chapters of this dissertation. The main rationale is that SI training affects different executive functions since SI implicates both domain general-executive functions and language-specific functions (see e.g. Hervais-Adelman & Babcock (2020), Abutalebi & Green (2007; 2008), Green & Abutalebi (2013)) to a more extreme degree than other forms of bilingualism (see Hervais-Adelman et al. (2017), Hervais-Adelman et al. (2015a), Elmer et al. (2014)), owed to performing overlapping processes of language reception, translation and production under time pressure (see Frauenfelder & Schriefers (1997)) and simultaneously highly active representations from both languages (see Paradis (1994), Dijkstra & van Heuven (1998)).

Regarding the **competence data**, the t-test results from analysing the F1 competence scores from the pre- and post- questionnaire show a significant difference between pre- and post test and indicate an improvement in SI competence which is concurrent with other studies that have found gains in interpreting performance after interpreting training (see Tzou et al. (2012), Hervais-Adelman et al. (2015a) and Yu & Dong (2022)). Results from a linear mixed model with F1 as dependent variable and Session as predictor and Participant as random effect only showed a trend toward significance and increased competence after training (see Figure 7). This may be rooted in the observation that two participants decreased their f1-score at post-testing (see Figure 6). Although the three analysed exam data are only anecdotal and cannot be used to analyse behaviour in the whole sample, they led to an important conclusion: I initially considered annotated exam evaluation sheets to be an objective and reliable measure, however, the three examples have shown that this way of evaluating competence might be far from reliable. One participant did not pass the final exam, and one participant performed worse on the final exam than in the initial exam. This however does not mean, that these students did not acquire or enhance their SI skill over the course of their Masters programme. Rather, the way the exam evaluation sheets evaluate competence or performance is inadequate from an instructional point of view. A performance graded with “good” at the end of the second semester of SI training does likely not correspond to a performance graded with “good” in the final interpreting exam at the end of the Masters programme since the demands placed on the SI students increase during the master's program. A performance graded “satisfactory” at the end of the SI training compared to a performance graded “good” at the beginning of training can still reflect an increase in SI skill and competence. It is highly probable that all participants did experience an increase in SI skill and ultimately competence, however, the exam evaluation sheets do not reflect this. Correspondingly, it is still possible that the two participants have increased their interpreting competence over training but were not able to achieve the advanced learning objectives underlying level 4 of the interpreting curriculum. Unfortunately, the analysed exam protocols do not allow this granularity of analysis. As argued by Beaton-Thome (2018), replacing the final exam with a more continuous assessment of competence would make it possible to assess a broader range of different aspects of competence (Beaton-Thome 2018: 158). In the end, an exam is also merely a snapshot in the performance of the interpreting student. Combining this proposed continuous assessment of SI competence with replacing grades by a more skill performance oriented approach could benefit a more granular assessment of SI competence and also boost the skill of the SI students to better reflect on their competence gains.

As Section 6.2 has shown, many researchers consent to the idea that after completing interpreting training, students can indeed have improved their interpreting competences (see e.g. (see Tzou et al. (2012), Hervais-Adelman et al. (2015a), Yu & Dong (2022))). However, this still must be distinguished from expert performance, that can only occur after years of consistently training interpreting skills on the job (see Tiselius & Hild (2017), Hoffman (1997), Kutz (2010)). The fact that two students showed worse performance in their final exam where they should have excelled assuming increasing interpreting competence over the course of training and performed better in their first exam, might indicate that the student interpreter’s performance still depends on some external factors such as stress. With increasing expertise, the trainee interpreters should be less affected by external factors such as stress (see Chabasse (2009) and Hoffman (1997)) and show a steadier performance over various interpreting jobs. The diverging results also highlight that not all students progress in the same manner within the same timeframe and their competence progress might differ individually.

Regarding the interpreting competence acquisition models, the significant difference between pre- and post-questionnaire TICQ-scores in the t-test and the marginally significant results of the linear mixed model that both showed an increase in interpreting competence allows the assumption that the participants' interpreting competence has shifted from the "knowledge about translation phase" to the "ability to act phase" as described by Kutz (2010). Respectively, participants should have moved from the "formation phase" over to the "controlled execution phase" as described by Chein & Schneider (2012)). Of course, the method does not allow for the observation of changes in brain anatomy or activation. The expansion-partial-renormalization hypothesis (see Lövdén et al. (2013)) and the dynamic restructuring model (see Pliatsikas (2020)) can therefore not be discussed with the data at hand.

Regarding the behavioural between-group comparison at baseline for the **Flanker task**, the main effects of Group in both RT and accuracy did not show a difference between groups at baseline (see Figure 8 and Figure 10), which corresponds to previous studies that did not find a difference between groups in inhibitory performance at baseline testing such as in Babcock et al. (2017), Dong & Liu (2016) and Rosier et al. (2019) and therefore also corresponds with the hypothesis formed on the basis of said literature. The significant main effect of Congruency on RT and accuracy, showing faster RT and higher accuracy for congruent than for incongruent trials (see Figure 8 and Figure 10) also corresponds to the hypothesis formed based on relevant Flanker task literature (see Verreyt et al. (2016), Luk et al. (2010), Dong & Xie (2014), van der Linden et al. (2018)). These behavioural findings were partly confirmed at electrophysiological level, since no significant difference in voltage in the N200 time-window between groups at baseline was found (see Figure 48). This again corresponds to the initially formed hypothesis based on previous studies that did not find a behavioural difference between groups at baseline in tasks that target the inhibitory executive functions (see Dong & Liu (2016), Babcock et al. (2017), Rosiers et al. (2019)). However, there was also no significant difference between Conditions (see Figure 48) which contradicts the hypothesis based on N200 literature that assumed more negative voltage for the incongruent condition in the N200 time-window (see Wild-Wall et al. (2008), Johnstone et al. (2009), Kousaie & Phillips (2012), Grundy et al. (2017b)).

The longitudinal behavioural results that RT in the Flanker task sped up over testing (see Figure 13) correspond to previous results of Babcock et al. (2017) who observed faster reaction times in the monitoring section of the Attention Network Task after two years of SI training. However, in their study, this effect was not SI specific, since all tested participant groups showed faster RT at the post-test compared to the pre-test. Similarly, Dong & Liu (2016) who used a Stroop task found faster RT across all groups after one semester of respective interpreting, translation or language training. This effect was therefore also not SI specific. Since the Control group in this study was only tested at baseline, no definite statement can be made regarding whether the faster reactions times were interpreting specific or also occurred in the translation students. The finding that Congruency had a significant effect on reaction time with the incongruent trials showing longer reaction times (see Figure 13) are again in line with the initially formed hypothesis that was based on previous findings literature (see Verreyt et al. (2016), Luk et al. (2010), Dong & Xie (2014), van der Linden et al. (2018)). Furthermore, the present behavioural results show that participants did increase their overall accuracy indicated by the significant effect of Session 4 and Session 5 on accuracy (see Figure 16), an effect not observed in the longitudinal studies by Dong & Liu (2016) and Babcock et al. (2017). However, the significant effect of Session 4 on accuracy did not mark an increase but a significant decrease in accuracy (see Figure 16), an example which showes nicely how important longitudinal designs can be that go beyond pre- and post-tests, offering a more detailed look into performance progression in interference suppression during SI training. Four participants (P01, P06, P08 and P10) attended their Erasmus-semester in Semester 3 (Session 4) and were not able to practice a lot of interpreting at this point of testing which might explain the drastic decrease in accuracy at this testing point.

The longitudinal behavioural results could not be statistically confirmed by the electrophysiological results. ERP-results in the N200 time-window showed that there was no effect of Session on Voltage across the time of testing (see Figure 52) which corresponds to previous behavioural results by van de Putte et al. (2018), who did not find a behavioural improvement in inhibitory control after 9 months of training. The effect of Condition on Voltage was not significant either, which again contradicts the behavioural findings that showed slower RT and decreased accuracy for incongruent trials as well as relevant Flanker task literature that associates more larger (more negative) voltage in the N200 time-window for the incongruent condition (see by Kousaie & Phillips (2012), Wild-Wall et al. (2008), Johnstone et al. (2009) or Grundy et al. (2017b)). Only the findings of significantly higher latencies for the incongruent than for the congruent condition confirmed the literature that the incongruent condition is more effortful and required more interference suppression (see Figure 55). As can be seen in Figure 51, voltage in both conditions does change across Sessions, decreasing until Session 3 and then slightly increasing again which is partly statistically confirmed by a significant interaction of Session 3 and condition (see Figure 52). Lower voltage in the N200 time window has been associated with increased resource allocation towards inhibitory control (see Cespón & Carreiras (2020)), so according to Figure 51, participants seemed to allocate more resources towards interference suppression until Session 3, after which they needed fewer resources. This might reflect that participants actively need their interference suppression to modulate the activation of words from both language system during SI (see Paradis 1994). However, since the ERP results could not be confirmed statistically, interpretation of the raw data is not reliable. Furthermore, behavioural results showed that accuracy significantly dropped at Session 4 (see Figure 16), so the increased voltage in the N200 time-window is probably rather associated with that than with a decreased need to engage inhibitory control due to increased SI competence. At this point it needs to be mentioned that inhibition is the executive function that has been shown to be least affected by SI training; an SI advantage in inhibitory controlled was rarely found (see Nour et al. 2020 and Hu & Fan 2021) and the possibility remains that inhibition is just not as affected by managing multiple languages in the brain as other executive functions (Nour et al. 2020).

Lastly, the observation that voltage hits a negative peak at Session 3 is mirrored by the frequency analysis of P01. Especially in the incongruent condition in P01, alpha power increases until Session 3, after which is decreases for Session 4 only to increase again at Session 5 (see Figure 87). Increased alpha power has been associated with more inhibition (Klimesch et al. 2007). Although this is of course an anecdotal observation from one participant, it mirrors the course of voltage progression of the whole population.

For the **colour-shape switch task**, the result that there are no differences between groups in reaction times at baseline (see Figure 18) correspond to results by Babcock et al. (2017), Dong & Liu (2016) and Rosiers et al. (2019) who all applied the colour-shape switch task and did not find a difference in RT between groups at baseline testing. The finding of a main effect of Congruency on RT with switch trials having longer RT and lower accuracy than repeat trials (see Figure 18) corresponds to relevant task switching literature (see e.g. Garbin et al. (2010), Baene et al. (2015), Prior & Gollan (2011), Declerck et al. (2017)). However, the result that participants do differ in accuracy at baseline (see Figure 20) but that this difference is due to the interpreting group making more errors than the control group does contradicts the findings by Rosiers et al. (2019) and Babcock et al. (2017) who did not find a difference in accuracy in the colour-shape switch task between groups. However, since the effect was driven by lower accuracy in the experimental group, no inherent advantage of the SI group at baseline is given, which in turn confirms my initial hypothesis that was formed on said literature. This pattern could be observed in the between-group comparison at baseline of the electrophysiological results for both the P300 component and the N200: The finding that there was a significant effect of Group with the Control group showing significantly lower P300 voltages than the Experimental group (see Figure 57) contradicts previous studies that did not find a behavioural difference between groups at baseline in the colour-shape switch task (Babcock et al. (2017), Dong & Liu (2016) and Rosiers et al. (2019)). According to relevant ERP-literature (see e.g. Daffner et al. (2010) and Polich (2007)), lower voltage in the P300 time-window (a smaller and more negative P300 amplitude) corresponds to increased task demands (see Daffner et al. (2011) and Cespón & Carreiras (2020)), increased attentional resource allocation to attentional shifting (see Polich (2007) and increased working memory load (Morrison & Taler (2020)). It therefore seems like the Control group allocated more attentional resources to switching attention and updating their working memory. The same goes for the N200 time-window, where the control group displayed significantly lower voltages than the Experimental group (see Figure 65). Together with the P300 results, this indicates more resource allocation to attention switching, stimulus categorization and working memory updating (see Daffner et al. (2010) and Polich (2007), Cespón & Carreiras (2020) Morrison & Taler (2020)) and actually speaks for a pre-existing advantage of the SI group. If the Experimental group also showed equal or better behavioural performance at baseline, this would indicate that the SI group possessed an inherent advantage in allocating attention, interference suppression, stimulus categorization and working memory updating. However, the behavioural analysis has shown that the SI group made significantly more errors at baseline (see Figure 20), indicating that the SI group likely performed at the cost of accuracy and that the seemingly lower task demands, resource allocation to attentional switching and stimulus categorization do not indicate an inherent cognitive advantage of the Interpreting group over the Control group.

The finding that there was no effect of Condition on voltage in the between-participant baseline analysis for the P300 time-window (see Figure 57) or for the N200-time window (see Figure 65) contradicts relevant ERP task switching literature (Timmer et al. (2017), López Zunini et al. (2019), Declerck et al. (2021), López Zunini et al. (2019), Periáñez & Barceló (2009) or Chen et al. (2022), Cespón & Carreiras (2020)) but speaks for the conclusion that both groups performed quite similarly, and the SI group did not show an inherent advantage in task switching.

In terms of the longitudinal behavioural results, the findings that participants significantly decreased their overall RT across sessions until the end (see Figure 23) is consistent with the results by Dong & Liu (2016) and Babcock et al. (2017), who both used the colour-shape switch task and found that participants had become faster after their respective training. However, as already mentioned in the Flanker task section, this effect in their study occurred for all participant groups and was therefore not SI specific. The linear mixed model for the switching costs did not reveal a significant effect of Session on switching costs (see Figure 26). This contradicts the findings by Dong & Liu (2016), who found that the interpreting groups showed significant gains in switch costs, therefore decreasing their switch costs significantly. This could be due to a lack of statistical power in the present study, since many more SI participants (57) participated in Dong & Liu's (2016) study. The present finding that accuracy increased across testing (see Figure 27) contradicts the results by Dong & Liu (2016) and Babcock et al. (2017), who did not find an increase in accuracy at the post-test. Furthermore, the significant Session:Condition interaction for Session 5 in the present data (see Figure 30) showed that accuracy decreased in Session 5 for the switch condition while it increased for the repeat condition. Moreover, due to the repeated testing in this longitudinal design, modulations in accuracy than can be seen in Session 3 that were not evident in other longitudinal studies only using pre- and post-tests could be revealed (see Figure 28). The findings that switch trials showed significantly slower reaction times and lower accuracy than repeat trials (see Figure 23 and Figure 29) correspond to relevant literature on the colour-shape switch task literature (see e.g. Garbin et al. (2010), Baene et al. (2015), Prior & Gollan (2011), Declerck et al. (2017)).

With regard to the longitudinal electrophysiological data from the P300 time-window, the finding that the effect of Session 5 on voltage did not reveal a significant result (see Figure 61) contradicts the findings from the behavioural data as well as previous behavioural studies that found a improvement in task switching performance at the end of a longitudinal testing design (see Dong & Liu (2016) and Babcock et al. (2017)). However, Session 3 and Session 4 revealed significant effects on voltage and it can be seen in Figure 60 that voltage increased in both conditions in Sessions 3 and 4. According to relevant P300 literature, increased voltage in the P300 time-window (more positive or larger P300 amplitudes) is associated with less resource allocation to attentional switching, stimulus categorization and working memory updating (see (see e.g. Daffner et al. (2010), Polich (2007), Morrison & Taler (2020), Cespón & Carreiras (2020)). Together with the behavioural findings that reaction times became faster (see Figure 23), and accuracy increased overall over the course of testing (see Figure 29 and Figure 27), this indicates an improvement in attentional resource allocation during task switching. The significant main effect of

Condition on Voltage with the repeat condition showing more positive voltage than the switch condition (see Figure 61) corresponds to relevant P300 task switching literature (Lopez Zunini (2019), Chen et al. (2022), Perianez Barcelo (2009), Cespon Carreiras (2020)) and also to the longitudinal behavioural findings that switch trials elicited longer reaction times and lower accuracy. Furthermore, the switch condition also showed significantly longer latencies (see Figure 64) that, together with the lower voltage, correspond to previous literature that associates switch trials with longer P300 latencies than repeat trials (see Céspon Carreiras (2020)). The increase in voltage in the P300 time-window until Session 4 and the following decrease in Session 5 (see Figure 60) is not mirrored in the overall accuracy but in the accuracy data per condition that shows a decrease in accuracy at Session 5 but only for the switch condition. However, accuracy is still significantly higher at Session 5 for both conditions than at baseline. It is possible that participants have reached a plateau where their cognitive ability to improve attention cannot be further boosted by SI training, which is reflected in the decrease in voltage in the P300 time-window in Session 5. When looking at the overall accuracy that increases until the end of testing at Session 5 (see Figure 29 and Figure 27) and the reaction times data that also improved until Session 5 (see Figure 23), the decrease in P300 voltage at Session 5 might indicate an automation of processes, so participants need to allocate fewer cognitive resources to attention switching while still keeping up performance. A plateau of this kind has been speculated about by Dong & Liu (2016) who theorized that there might be a development curve of executive function enhancement following SI. The mechanisms of this plateau are rooted in models such as by Chein & Schneider (2012) and Paap (2018). Chein & Schneider (2012) have suggested that after L2 learners transition from a phase the requires high amounts of executive control to a phase where they perform less automatically and rely less on executive control while still performing well. Formulated on this basis, the controlled dose hypothesis by Paap (2018) suggests that any activity that requires a lot of executive functions initially will boost executive functions at first but this boost will fade with the transition from controlled to more automatic processing (Paap 2018: 454). Transferred to the interpreter advantage and the current ERP results for attentional switching in the colour-shape switch task, this would suggest that the ability to reallocate attention during task switching was initially boosted by progressing SI training and gains in competence but then reached a plateau and shifted to more automated and efficient processing.

Results from the N200 time-window and the associated function of interference suppression are far less revealing: The finding that no effect of Session on voltage was significant across sessions (see Figure 69) does contradict the hypothesis formed based on literature that has shown an improvement in behavioural performance in task switching (Babcock et al. (2017) and Dong & Liu (2016)) but corresponds to previous findings by van de Putte et al. (2018) who did not find an improvement in inhibitory executive control after SI training, the N200 component in task switching representing the interference suppression aspect of task switching (Jamadar et al. (2015), Cespon & Carreiras (2020)). However, it is in line with the results of the Flanker task that, while modulations of the N200 voltage were visible in the raw data, they did not show a significant effect of any Session on voltage. Modulations in the raw data in the N200-time window were also visible for the colour-shape switch task, with voltage increasing until Session 3 followed by a decrease in Session 4 (see Figure 68). According to literature examining the N200 component, lower (more negative) N200 amplitudes are associated with more resource allocation to interference suppression (see Cespón & Carreiras (2020)). The raw data in Figure 68 therefore theoretically show that participants needed to allocate less interference suppression during task switching until Session 3 and needed to increase the resource allocation to interference suppression again slightly in Session 4. However, this trend is not statistically confirmed and therefore to be considered with caution.

A linear mixed model with switch costs per participant as dependent variable, Session as predictor and Participant as random effect showed a significant effect of Session 4 and Session 5 on switching costs. This mirrors the progression of voltage in both conditions in Figure 68 where the switch condition becomes more negative again especially in Session 4. The finding for significant switch costs in Session 4 and 5 that become more positive does not correspond to previous literature that has shown gains in switch costs after SI training (see Dong & Liu (2016) and Babcock et al. (2017)). Lastly, the finding that Session did not affect N200 latency significantly does not correspond to the expectations formed based on N200 literature that assigns later latencies with increased cognitive resource allocation (Jamadar et al. 2015), however, the finding of a significant effect of Condition on latency with longer latencies for switch condition than for the repeat condition corresponds to relevant N200 literature (Cespón & Carreiras 2020).

The results that there is no significant difference in RT between groups at baseline for the **2-back task** (see Figure 31) correspond to results by Babcock et al. (2017) who did not find a difference between groups at baseline in two working memory measures and to results by Rosiers et al. (2019) and Dong & Liu (2016) who did not find a difference in RT between groups for the 2-back task at baseline testing. The result that accuracy was significantly higher for the Experimental group than for the control group at baseline (see Figure 32) contradicts the results by Rosiers et al. (2019) who did not find a difference between groups at baseline for accuracy in the 2-back task. Since there was no significant difference in reaction time, it might be possible that the Control group traded accuracy for speed. This result speaks for an inherent advantage in working-memory updating for the SI group which is in line with the possibility that individuals who already possess some cognitive predisposition enter the MA Conference Interpreting (see Babcock et al. (2017)). The finding that the no-match condition was characterized by significantly higher accuracy (see Figure 32) corresponds to relevant n-back task literature such as by van der Linden et al. (2018), Szmalec et al. (2011) and Barker & Bialystok (2019). For the ERP data, the result that there is no significant difference between groups at baseline in the P300 time-window (see Figure 73) corresponds to the present behavioural results as well as to previous literature of behavioural studies which did not find a difference in WM-updating performance between groups at baseline (see e.g. Dong & Liu (2016), Rosiers et al. (2019) and Babcock et al. (2017)). The result that Condition had a significant effect on Voltage with the no-match condition showing lower voltage than the match condition (see Figure 73) corresponds to the present behavioural findings that associated the no-match condition with longer RT as well as the results in relevant n-back literature (see Barker & Bialystok (2020)).

For the experimental group, the progression of reaction times across sessions revealed significant main effects of Session on RT for Session 3, 4 and 5 and therefore showed that RT decreases steadily until the end of training (see Figure 34). This is in line with the results by Dong & Liu (2016) who found that the SI group and not the bilingual control group improved their reaction times after training. However, the translator group in their study also improved their reaction times after training. Since the translator control group in the present study was not tested longitudinally, it cannot be said for certain that the improvement in RT across sessions is interpreting specific. The result that accuracy also increased over the course of testing, especially the significant increase in both condition in Session 4 and 5 (see Figure 37) corresponds to findings by Babcock et al. (2017) who found a numerical decrease of errors for the SI group and the control group. Figure 36 shows that raw data accuracy in the match condition decreased slightly, however, this was not supported by a significant interaction between Session 5 and Condition (see Figure 37).

The longitudinal behavioural findings could not be confirmed statistically by the ERP data: The result that none of the Sessions had a significant effect on Voltage in the P300 time-window across all Sessions (see Figure 77) contradicts the present behavioural literature as well as previous literature that has found a significant effect of post-test on reaction time or accuracy in working memory measures (see Dong & Liu (2016) and Babcock et al. (2017)). The raw data in Figure 76 depicts that there were modulations of voltage in the P300 time-window and, despite an initial decrease in voltage, voltage in both conditions increased until Session 4, followed by a slight decrease. Higher voltage in the P300 time-window has been associated with less resource allocation to working-memory updating through stimulus categorization (see Polich (2007), Cespón & Carreiras (2020)). Together with the significant improvements in RT and accuracy until the end of testing, this speaks in favour of an improvement of WM-updating ability across SI training. Furthermore, since voltage decreased at Session 5, but RT and overall accuracy remained stable, this might indicate a plateau where SI training could not further cognitively boost executive functions. However, since the ERP data is not supported by statistics, interpretation in this direction must be handled with caution. Lastly, the finding that Session and Condition did not show a significant effect on latency did not correspond to existing n-back literature such as by Barker & Bialystok (2019), who found longer P300 latencies on match than on no-match trials for the 1-back task.

The behavioural results of the between-group comparison at baseline of the **3-back** task did not show a significant difference between groups at baseline in reaction times or accuracy (see Figure 38 and Figure 39) and therefore correspond to the findings by Dong & Liu (2016) and Rosiers et al. (2019) who used a 2-back task but did not find a significant difference between groups at baseline and to the findings by Babcock et al. (2017) who did not find a difference between groups in two different working memory measures. These findings are especially interesting in regard to the behavioural 2-back results at baseline that indicated an inherent advantage in WM-updating for the SI group. The lack of behavioural difference between groups in the 3-back tasks shows that when processing effort on working memory was higher than in the 2-back task, no advantage of the SI group could be detected, confirming the initial hypothesis based on literature by Dong & Liu (2016), Rosiers et al. (2019) and Babcock et al. (2017). The finding that Condition had a significant effect on accuracy with lower accuracy for the match than for no-match condition (see Figure 39) corresponds to relevant n-back literature (see van der Linden et al. (2018)).

These behavioural findings at baseline were mirrored by the ERP-data that showed that there is no significant difference between groups at baseline in the P300 time-window (see Figure 79). These findings again correspond to the initial hypothesis formed on the basis of previous literature of behavioural studies that did not find a difference in WM-updating performance between groups at baseline (see e.g. Dong & Liu (2016), Rosiers et al. (2019) and Babcock et al. (2017)). Furthermore, the result that Condition had a significant effect on Voltage with the no-match condition showing lower voltage than the match condition (see Figure 79) corresponds to the hypothesis formed on the basis of results from relevant n-back literature (see Barker & Bialystok (2020)). In summary, neither behavioural results nor electrophysiological data suggest an inherent advantage of the SI group.

For the progression of reaction time, Session 3, 4 and 5 showed a significant main effect on reaction time and it became apparent that that RT decreased steadily until the end of training (see Figure 41). This again corresponds to the results from the 2-back task and to previous results by Dong & Liu (2016) who found that the SI group and not the bilingual control group improved their reaction times after training. The result that accuracy also increased over the course of testing (see Figure 42),

especially the significant increase in both condition in Session 4 and 5 corresponds to findings by Babcock et al. (2017) who found a numerical decrease of errors for the SI group and the control group. The highly significant effect of Condition on Accuracy with the match condition showing lower accuracy corresponds to relevant n-back task literature (see by van der Linden et al. (2018) and Szmalec et al. (2011)).
Similar to the longitudinal ERP results of the 2-back task, the electrophysiological results for the 3-back task did not statistically confirm the progress seen in the behavioural data: there was no effect of Session on Voltage in the P300 time-window across Sessions (see Figure 83) which contradicts the initially formed hypothesis based on previous literature that has found a significant effect of post-test on reaction time or accuracy in working memory measures (see Dong & Liu (2016) and Babcock et al. (2017)). Taking a look at the raw data in Figure 82 indicates that there are modulations in voltage in the P300-time window, showing e.g. that voltage in both conditions reached a positive peak in Session 3, decreased for Session 4 and slightly increased again in Session 5. Higher voltage in the P300 time-window has been associated with less resource allocation to working memory updating via stimulus categorization (see Polich (2007), Cespón & Carreiras (2020), Daffner et al. (2011)), so the observed trend would indicate that participants needed less cognitive resources for WM-updating in Session 3, more in Session 4 (maybe due to lack of training after the Erasmus semester) and less cognitive resources again after completing their four-semester SI training. However, since the ERP data could not be statistically supported, the interpretation of these trends must be considered cautiously. Lastly, the significant effect of Condition with lower voltage for the no-match condition is however in line with previous n-back literature (see Barker & Bialystok (2019)).

In the direct **comparison of the 2-back and 3-back task** in the SI group, results showed that participants performed faster and more accurately across the course of testing in both tasks (see Figure 44 and Figure 46), mirroring the separate reaction time and accuracy results reported previously. More importantly, the effect of Task on reaction time and on accuracy was significant, with the 3-back task always showing significantly slower RT and lower accuracy compared to the 2-back task (see Figure

44 and Figure 46) which corresponds to relevant n-back literature (Jaeggi et al. 2010), explaining that processing load can be varied with increasing N. This means that the 3-back task in general requires more cognitive processing than the 2-back task and, more specifically, different amounts of cognitive resource allocation to inhibition and working memory updating. The findings furthermore correspond to findings such as by Morales et al. (2015) that found that the 2-back task was easier to perform than the 3-back task. Mirroring these findings, the ERP data showed a marginally significant effect of Task on Voltage with the 3-back task showing lower (more negative) voltage in the P300 time-window (see Figure 85) which has been associated with increased cognitive resource allocation to WM-updating through stimulus-categorization (see Kok (2001), Cespón & Carreiras (2020), Daffner et al. (2011), Morrison & Taler (2020), Dong et al. (2015)). However, contrary to the behavioural effect of Session on Voltage, ERP results did not show a significant effect of Session on voltage (see Figure 85) which in turn mirrors the results from the separate 2-back and 3-back ERP analysis. Unfortunately, there were no significant 2-way or 3-way interactions that would have warranted an interpretation beyond the main effects (see Figure 85).

To sum up, the initial hypothesis formed in Chapter 7 that interpreters do not possess an inherent advantage in executive function compared to translators before they start training (see Dong & Liu (2016), Babcock & Vallesi (2017), Rosiers et al. (2019)) could be confirmed behaviourally and electrophysiologically with the exception of an inherent SI advantage at baseline for the WM-updating that surfaced in the 2-back task but disappeared however when memory load was increased in the 3-back task. The most promising result that could also be supported by statistically significant results was the improvement of the attention switching ability in task switching in the colour-shape switch task over the course of testing. After considering a performance at the cost of decreased accuracy for the SI group, an initial interpreting advantage in task switching of the SI group over the translation control group could be excluded which corresponded to the initial hypothesis formed on the basis of relevant literature (Dong & Liu (2016), Babcock et al. (2017), Rosiers et al. (2019)). Together with the improvement in SI competence scores from pre-training to post-

training indexed by increased TICQ-scores, both longitudinal behavioural data as well as electrophysiological data showed statistically significant evidence for an improvement of the attention allocation mechanism during task switching. Most importantly, the ERP-data could uncover a potential automation or cognitive performance plateau at Session 4, since voltage in the P300 time-window peaked at Session 4 while accuracy and reaction time remained steady or further improved for Session 5. As discussed above, a mechanism like this has been hypothesized in the context of bilingualism (see Chein & Schneider (2012), Pliatsikas (2020) and Paap (2018)) as well as SI (see Dong & Liu (2016)) but has so far not been shown in the context of SI and executive functions since pre- and post-test designs are not suited to uncover what happens between the two testing sessions.
Since electrophysiological results from the Flanker task, the interference suppression component of the colour-shape switch task and the 2-back and 3-back task were not significant and could not support the significant behavioural results, it is not possible to say for certain that cognitive improvements in interference suppression and working-memory updating exist beyond test-re-test effects. This is especially likely for interference suppression, for which effects of SI training has been shown the least in previous literature (see Nour et al. (2020)). As has become apparent in Figure 56 and Figure 72, participant variability was quite high which might be one cause of missing effects of Session.

Nevertheless, applying event-related potentials and especially applying a longitudinal design that goes beyond the pre- and post-test design was extremely illuminating: If I had only done a pre- and post-test, therefore only comparing Session 5 to Session 1 (baseline)[5], an extremely reduced picture would present itself: Reaction time and accuracy analysis would indicate that SI students had increased their reaction times and accuracy in all executive function tasks and would therefore have increased their ability for interference suppression, attentional shifting as well as working memory updating and stimulus categorization. However, in the pre- and post-test approach, ERP data would reveal simply no significant difference from baseline to Session 5

---

[5] Linear mixed models use Session 1 as reference and automatically compare all following sessions to Session 1 (baseline)

even though significant difference had surfaced at other testing points in-between. As the results have shown, this pre- and post approach would have been especially detrimental for the electrophysiological insights into attention shifting aspect indexed by the voltage within the P300 time-window, where voltage almost decreased back to baseline while the behavioural results all improved or remained steady, indicating more efficient attention shifting after Session 4. My more detailed longitudinal design revealed significant effects of Session 3 and 4, depicting a bridge-shaped progression of voltage over the course of testing and showing neuroscientific evidence that participants increased their attention shifting function over the course of testing. This shows neatly how important information about the neural mechanisms would have been lost in a pre- and post-test design using only behavioural data. Other studies applying a pre- and post-training design might be affected by this issue. For example Babcock et al. (2017) conducted a longitudinal study in a pre- and post-test design where they compared MA students of conference interpreting to MA translation students and students with non-language related degrees in inhibition, attention, switching and short-term memory. They did not find that interpreters outperformed the other groups post-training in any executive function after a 2-year training period, they only found gains in short-term memory. Similarly, van de Putte et al. (2018) compared the performance as well as neural activity during the Simon task, a colour-shape switch task and a verbal fluency task of interpreting students and translation students after a 9-month training period. They did not find any behavioural differences between groups after the training period but found differences between groups in brain activity during the Simon and the colour-shape switch task. The lack of behavioural differences might stem from the assumption that reaction times are not fine-grained enough, however, it is also possible that there were differences between groups in the middle of testing and they have renormalized at the end of training. Since the authors also found increased connectivity in the fronto-basal ganglia subnetwork and a network of the cerebellum and the supplementary motor area, both networks being associated with executive functions and language control, the SI group might have become more efficient. Having reported these examples, the study at hand contributes to the research landscape of executive functions in simultaneous interpreting students by highlighting that testing at smaller intervals can benefit the understanding of what happens to executive functions or even structural and functional characteristics of the brain during training.

It has become apparent that executive functions do not behave in unison and individual differences play a big role in how executive functions progress. This has become most apparent in the N200-time window in the Flanker task and in the colour-shape switch task that both target interference suppression but seem to tap into different mechanisms of interference suppression, nevertheless. Although it is the same executive function (suppressing interference from the irrelevant Flankers or task rule) it behaves differently. Half of the participants in Session 4 participated in an Erasmus semester abroad and did not participate in SI training. The is likely reflected in the low accuracy in Session 4 of the Flanker task. Interestingly, voltage in the N200 time-window reached a positive peak at Session 4. However, it is likely that this positive peak is not associated with more ease in task performance and more efficiency in executive functions but rather in not using interference suppression as much which resulted in lower accuracy. Voltage then decreased again for Session 5, but not below the second positive peak at Session 2. In the colour-shape switch task, the Erasmus semester also seemed to have an impact. Although accuracy reaches a positive peak at Session 4, voltage reaches a negative peak, likely reflecting more involvement of interference suppression. It seems like the need to keep up performance after the Erasmus semester warranted the increased engagement of interference suppression, since no interpreting was done during the semester abroad. Interestingly, the Erasmus semester seemingly did not affect the aspect of attentional control and working memory updating or stimulus-response mapping, since neither accuracy nor voltage decreased at Session 4 in the P300 time-window in the colour-shape switch task or the 2-back or the 3-back task. Pausing SI training only seemed to affect interference suppression.

The trends that could be observed in terms of the progression of different executive functions have different implications for SI didactics. On the one hand, it has become clear that abilities in all executive functions showed a steep increase after the beginning of training, reflecting the results from the study by Dong & Liu (2016) that has shown an improvement in behavioural performance in different executive functions after only one semester of training. On the other hand, it could be observed that although there was a trend for the experimental group to progress in a certain way, the individual data showed a large heterogeneity when students reached a high or low peak in their respective executive functions (see in Figure 56 and Figure 72). This emphasizes that individuals do respond to the increased control demands at a differential pace and therefore shows again that, since all brains are different, the impact of SI training and the associated competence acquisition is therefore also quite individual, meaning that SI training might not be a programme that universally fits everyone. There is no definitive evidence in the results that the Erasmus semester was the cause of a setback in different executive functions. However, it could be argued that one semester of a whole years of absence from SI training might have an impact on the training progress and that students might need a longer time to get back into the SI practice. Such hypotheses need to be examined in carefully controlled future studies.

In terms of the implications of the results of this dissertation for SI competence acquisition, I would like to refer back to the assumption that becoming more adept at managing the increased cognitive control demands of SI and therefore increasing SI competence leads to a more efficient use of executive functions, also named the interpreter advantage (see García (2014)). However, as discussed earlier, the apporach of using the graded exam evaluation sheets to capture gains in SI competence did not prove as fruitful since an alleged decrease in graded performance at the end of SI training compared to the beginning of SI training cannot be indicative of a lack of competence acquisition. This cross-sectional approach of assessing performance and SI competence makes it difficult to gauge the degree of competence acquisition. A proposition for didactics could be to offer a closer-knit and at the same time more diverse assessment of the SI competence in the final semester. Beaton-Thome (2018) has summarized several solution-oriented approaches. In the sense of situated learning within a community of practice, she suggests that constellations of experienced SI tutors and students from the beginning of training could be fruitful for a more integrated approach to SI expertise acquisition. Most importantly however, she suggests that the final interpreting exam at the end of training should be replaced with a more continuous assessment of competence (Beaton-Thome 2018: 157–158). This is in line with the approach that was taken in the empirical study in this dissertation to observe the progression of executive functions. Similar to the issues with executive function assessment, a more continuous assessment of SI competence would go beyond simple pre-training and post-training assessments of performance and competence. Beaton-Thome (2018) also emphasized that being in an all-or-nothing exam situation might not reflect the competence of students all too well, which corresponds neatly to my observation. She proposes that students would have the possibility to present different aspects of their acquired competences in more differentiated final assessments in the last semester of training (Beaton-Thome 2018: 158).

On a broader scale, the results obtained in the present study furthermore provide important insights for cognitive bilingualism research. For one, the sample examined in this study, simultaneous interpreters, who switch between languages frequently and therefore must at least initially allocate increased effort towards inhibiting the language representations currently not used for speaking under great time pressure, enriches the assumption that “good language switchers are good task switchers” (Prior & Gollan 2011). The findings contribute to the discussion within the bilingual advantage debate that research should move away from the yes or no binary and move towards emphasizing the moderating factors that affect executive functions in bilinguals. The current study supports the assumption by Woumans & Duyck (2015), that language switching practice is one of the leading components that affect the functionality and organization of the bilingual or multilingual brain. At this point however, the lack of longitudinal data from a translator or non-translating bilingual control group limits the scope to which definitive statements can be made.

On an electrophysiological level, bilingualism has been associated with increased resource allocation towards different aspects of cognitive control, e.g. monitoring or inhibition (see e.g. Markiewicz et al. 2023) presumably due to the increased control demands posed by managing multiple languages. Especially the results from the task switching paradigm in this study offer the perspective that with persistent and extrensive training of language control, the executive functions can let go more as processing becomes more automated. This also corresponds to the model proposed by Pliatsikas (2020) that explains how the bilingual brain restructures itself while gaining more and more control over a second language. The results from the task switching paradigm, together with the fronto-parietal electrodes that were chosen to examine the response of the brain to the stimuli requiring a task switch, are in line with the findings of Hervais-Adelman et al. (2011) who propose that a fronto-parietal network supports general switching mechanisms in bilingual language control. Ultimately, this study rooted in the setting of SI, provides further evidence that the domain general networks of language control are also implicated in bilingual language control and as an extension also play a crucial role in modeling SI neurocognitively as a form of extreme language control, as it was done by Hervais-Adelman & Babcock (2020).

So, to conclude, under some reserve discussed in the following chapter, the results have shown that executive functions do have the potential to change during SI training, although the present data could only show this in the case of attention allocation during task switching. This highlights the plasticity of the bilingual brain, its ability to reshape in reaction to different experience, one of the most complex being simultaneous interpreting.

10. Conclusion, limitations, and future research

This dissertation has positioned simultaneous interpreting as a cognitively demanding task that requires the coordination of overlapping processes of language reception, translation and production processes and therefore creates high demands for language control that also implicates executive functions. Simultaneous interpreters have therefore been assumed to possess enhanced executive functions as part of the interpreter advantage hypothesis. In the recent past, this assumption has been tested by means of neuroscientific methods that highlighted the involvement of executive functions during simultaneous interpreting and the structural and functional changes that occur in the brain as a function of interpreting training. However, when comparing simultaneous interpreters' performances in different executive functions to other bilingual populations such as translators or language teachers in cross-sectional studies, inconsistent results have been found. Cross-sectional studies have therefore been criticized to only measure performance at one point in time and, since interpreting competence and the associated cognitive performance is difficult to gauge by grouping participants by years on the job, longitudinal studies have been proposed to remedy this issue. However, the longitudinal studies that were conducted mostly relied on behavioural measures that risk obscuring the mechanisms of the brain by motor response latencies. Furthermore, longitudinal studies in this field always applied a pre- and post-test design to observe interpreting students' competences and executive function gains after training, concealing what happens during the process of competence acquisition. To mitigate these issues, I conducted a longitudinal study that goes beyond the pre- and post-test methodology, testing the performance in three executive functions tasks of a group of interpreting students continuously throughout their Master of Conference Interpreting. By using a triangulation of event-related potentials, behavioural measures as well as competence data, I was hoping to provide a more accurate and explicit insight into how executive functions are affected by training extreme language control through interpreting training. To see whether the SI students possess inherently enhanced executive functions since they must pass an aptitude test to enter their Master studies, I compared the performance of the SI students to a group of translation students at baseline.

Results from the TICQ-score revealed that self-rated interpreting competence increased from the baseline to the end of SI training. Results from the behavioural as well as ERP data showed that besides low processing effort working memory tasks, no inherent SI advantage could be seen at baseline. For the longitudinal data, only the attention reallocation function during task switching provided reliable results from the combination of behavioural and ERP data, describing that the ability to reallocate attention during task switching is enhanced during SI training. However, since behavioural results continued to increase or remained steady at a certain point before the end of training while ERP data showed a decrease in the engagement of resource allocation to attention switching, a cognitive performance plateau and more efficient processing could be assumed. Hypothesis concerning an improvement in interference suppression and working memory updating could not be statistically confirmed in a combination of behavioural and ERP-data.

Nevertheless, another key finding is that by applying a pre- and post-test design, important information on the neural processes regarding the executive function would have been lost, since there was no significant difference between voltage in Session 1 and Session 5 in any task, however, there were significant effects of Session 3 and 4, for example in the colour-shape switch task. This reinforces the importance of applying longitudinal neuroscientific methods over cross-sectional methods and more importantly longitudinal methods that go beyond pre- and post-tests. The combination of behavioural data, electrophysiological data and competence data successfully supported the attempt to construct a full picture of effect of SI training on executive functions. The differential results for the different executive functions has furthermore highlighted that executive functions are not easily all put in one box and need to be chosen carefully.

This study contributes to existing longitudinal research on SI training and executive functions and also encourages testing executive functions longitudinally in comparison with other modalities of bilingualism such as translators, language learners or even language learners with impaired executive functions such as individuals with ADHD. Moreover, the study has highlighted that SI competence acquisition has the potential to not progress linearly and that gaps in training such as semesters abroad may have an impact on executive functions supporting SI competence acquisition. Lastly, I would like to emphasize the importance of studying simultaneous interpreting as well as written translation to further the understanding of how different circumstance of language use changes the brain.

There are some important limitations to this study that absolutely must be addressed. First and foremost, it is of course not ideal to only compare the interpreting group to the translator group at baseline. Had the Covid19-pandemic not restricted human contact and therefore lab use, and also had it not drastically lowered the motivation of the participants to commit to a longitudinal study, it would have been ideal to collect participants from both groups from the beginning and test them longitudinally in parallel throughout their respective MA of Conference Interpreting and MA Translation programmes. By only comparing between groups at baseline, I was able to account for possible inherent advantages of the interpreting group but by continuing to the study only with the experimental group, I was not able to draw conclusions about the translator group and could only compare the progress of the participants to themselves. I was consequently only able to make assumptions about how executive functions are modulated by interpreting training but was not able to say with absolute certainty that it was the interpreting training that did the trick. Future studies should therefore absolutely compare translation students and interpreting students longitudinally in both behavioural as well as electrophysiological measures to secure the promising findings that were brought about by the study at hand. This way, test-retest effects can be ruled out with more certainty.

Furthermore, another limiting factor concerns the number of participants. Although it is not uncommon in neuroscientific studies to have low participant numbers, this is one crucial factor that decreases statistical power and limits the generalizability across studies. Furthermore, a higher number of participants might have allowed for the success of different statistical analysis methods such as sliding contrast analysis, for which my participant number was simply too low. The low participant number was also rooted in the fact that it is extremely difficult to locate individuals at our campus that have not had any prior experience with interpreting and therefore made the pool of possible participants extremely small. Furthermore, the EEG and furthermore a longitudinal study with the EEG-technique can be a difficult selling point, since adjusting the EEG cap and filling in the gel to lower impedance can be very tedious for participants. Despite a great effort at advertising the advantages of the EEG, not all individuals feel comfortable with neuroscientific methods and the prospect of committing to a longitudinal study. Future studies should nevertheless strive to find as many participants as possible to ensure adequate statistical power and a wide range of possibilities for statistical analysis.

One more limitation lies in the competence assessment. The TICQ is an excellent tool and an attempt to establish a gold-standard instrument for the assessment of translation and interpreting competence. As all other self-report tools however, it is not free of the commonly known self-report bias that might skew the results. Additionally, for future studies, it would be helpful to administer the TICQ at every testing point and not only as a pre- and post-test measure.

In addition to the proposition of remedying the participant number and the parallel testing of the control group, future research leaves many options for follow-up studies. It would be interesting for example to continue testing in the field of professional interpreters and translators to see which role longitudinal professional experience plays in executive functions. Of course, using different neuroscientific measures than the EEG, such as fMRI or PET could also provide promising results but may not be easily adapted.

Moreover, it could be interesting for future studies to compare different bilingual populations longitudinally such as written translation, second language learners or even how professional language training affects individuals with impaired executive functions such as individuals with ADHD.

Lastly, taking a step back from the longitudinal approach, it might also be interesting to investigate which short-term effect translation and interpreting may have on executive functions, as it is currently being examined by Habig et al. (forthcoming).

References

Abutalebi, Jubin, Pasquale A. Della Rosa, David W. Green, Mireia Hernandez, Paola Scifo, Roland Keim, Stefano F. Cappa & Albert Costa. 2012. Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral cortex (New York, N.Y. : 1991)* 22(9). 2076–2086.

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 W. 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 W. Green. 2016. Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition* 19(4). 689–698.

Ahrens, Barbara. 2011. Neurolinguistics and interpreting. In Yves Gambier & Luc van Doorslaer (eds.), *Handbook of Translation Studies*: *Volume 2*, 105–107. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Ahrens, Barbara. 2017. Interpretation and Cognition. In John W. Schwieter & Aline Ferreira (eds.), *The handbook of translation and cognition* (Blackwell Handbooks in Linguistics), 445–460. Hoboken, New Jersey: Wiley Blackwell.

Ahrens, Barbara, Eliza Kalderon, Christoph M. Krick & Wolfgang Reith. 2010. fMRI for exploring simultaneous interpreting. In Daniel Gile, Gyde Hansen & Nike K. Pokorn (eds.), *Why Translation Studies Matter*, 237–247. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Albl-Mikasa, Michaela. 2012. The importance of being not too earnest: a process- and experience based model of interpreter competence. In Barbara Ahrens, Michaela Albl-Mikasa & Claudia Sasse (eds.), *Dolmetschqualität in Präxis, Lehre und Forschung*: *Festschrift für Sylvia Kalina*, 59–92. Tübingen: Narr Francke Attempo.

Amos, R.M. & M.J. Pickering. forthcoming. A new perspective on models and theories of simultaneous interpreting. *Target: International Journal of Translation Studies*.

Andres, Dörte. 2013. Das Konzept Freitagskonferenz: Expertiseentwicklung durch berufsorientierte Lehre. In Silvia Hansen-Schirra & Don Kiraly (eds.), *Projekte und Projektionen in der translatorischen Kompetenzentwicklung* (61), 237–255. Frankfurt am Main: Peter Lang GmbH.

Ansaldo, Ana I. Ladan Ghazi-Saidi & Daniel Adrover-Roig. 2015. Interference Control In Elderly Bilinguals: Appearances Can Be Misleading. *Journal of clinical and experimental neuropsychology* 37(5). 455–470.

Antonenko, Pavlo, Fred Paas, Roland Grabner & Tamara van Gog. 2010. Using Electroencephalography to Measure Cognitive Load. *Educational Psychology Review* 22(4). 425–438.

Aparicio, Xavier, Karin Heidlmayr & Frédéric Isel. 2017. Inhibition Efficiency in Highly Proficient Bilinguals and Simultaneous Interpreters: Evidence from Language Switching and Stroop Tasks. *Journal of psycholinguistic research* 46(6). 1427–1451.

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.

Babcock, Laura E. 2015. *The neurocognitive fingerprint of simultaneous interpretation:* Ita; Trieste; Sissa.

Baddeley, A. 2000. The episodic buffer: a new component of working memory? *Trends in cognitive sciences* 4(11). 417–423.

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

Baddeley, Alan. 2010. Working memory. *Current biology : CB* 20(4). R136-40.

Baddeley, Alan D. 1983. Working memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 302(1110). 311–324.

Baddeley, Alan D. & Graham Hitch. 1974. Working Memory. In Gordon H. Bower (ed.), *Psychology of Learning and Motivation*: *Advances in Research and Theory* (Psychology of Learning and Motivation v.8), 47–89. Burlington: Elsevier.

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.

Bajo, Maria Teresa, Francisca Padilla & Presentación Padilla. 2000. Comprehension processes in simultaneous interpreting. In Andrew Chesterman, Natividad Gallardo San Salvador & Yves Gambier (eds.), *Translation in Context*: *Selected contributions from the EST Congress, Granada 1998* (39), 127–150. Amsterdam, Philadelphia: John Benjamins Pub. Co.

Banich, Marie T. & Rebecca J. Compton. 2018. *Cognitive neuroscience*. Cambridge: Cambridge University Press.

Barker, Ryan M. & Ellen Bialystok. 2019. Processing differences between monolingual and bilingual young adults on an emotion n-back task. *Brain and cognition* 134. 29–43.

Bastian, Claudia C. von, Alessandra S. Souza & Miriam Gade. 2016. No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of experimental psychology. General* 145(2). 246–258.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1).

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.

Berg, Esta A. 1948. A simple objective technique for measuring flexibility in thinking. *The Journal of general psychology* 39. 15–22.

Berken, Jonathan A. Xiaoqian Chai, Jen-Kai Chen, Vincent L. Gracco & Denise Klein. 2016. Effects of Early and Late Bilingualism on Resting-State Functional Connectivity. *Journal of Neuroscience* 36(4). 1165–1172.

Bernhardi, Rommy von, Jaime Eugenín & Kenneth J. Muller (eds.) (2017). *The plastic brain* (Advances in Experimental Medicine and Biology 1015). Cham: Springer International Publishing; Imprint: Springer.

Bialystok, Ellen. 2015. Bilingualism and the Development of Executive Function: The Role of Attention. *Child development perspectives* 9(2). 117–121.

Bialystok, Ellen. 2017. The bilingual adaptation: How minds accommodate experience. *Psychological bulletin* 143(3). 233–262.

Bialystok, Ellen, Fergus I. M. Craik, David W. Green & Tamar H. Gollan. 2009. Bilingual Minds. *Psychological science in the public interest : a journal of the American Psychological Society* 10(3). 89–129.

Bialystok, Ellen, Fergus I. M. Craik, Raymond Klein & Mythili Viswanathan. 2004. Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and aging* 19(2). 290–303.

Bialystok, Ellen, Michelle M. Martin & Mythili Viswanathan. 2005. Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism* 9(1). 103–119.

Bloomfield, Leonard. 1933 // 1984. *Language*, 1933rd edn. Chicago, London: University of Chicago Press; Univ. of Chicago Press.

Bokura, Hirokazu, Shuhei Yamaguchi & Shotai Kobayashi. 2001. Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clinical Neurophysiology* 112(12). 2224–2232.

Botvinick, M. M. T. S. Braver, D. M. Barch, C. S. Carter & J. D. Cohen. 2001. Conflict monitoring and cognitive control. *Psychological review* 108(3). 624–652.

Braver, Todd S. Jeremy R. Reynolds & David I. Donaldson. 2003. Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron* 39(4). 713–726.

Buelow, Melissa T. Bradley M. Okdie & Ashley B. Cooper. 2015. The influence of video games on executive functions in college students. *Computers in Human Behavior* 45. 228–234.

Bunge, Silvia A. Nicole M. Dudukovic, Moriah E. Thomason, Chandan J. Vaidya & John D.E. Gabrieli. 2002a. Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*(17;33(2)). 301–311.

Bunge, Silvia A. Eliot Hazeltine, Michael D. Scanlon, Allyson C. Rosen & J. D. E. Gabrieli. 2002b. Dissociable contributions of prefrontal and parietal cortices to response selection. *NeuroImage* 17(3). 1562–1571.

Burgaleta, Miguel, Ana Sanjuán, Noelia Ventura-Campos, Núria Sebastian-Galles & César Ávila. 2016. Bilingualism at the core of the brain. Structural differences between bilinguals and monolinguals revealed by subcortical shape analysis. *NeuroImage* 125. 437–445.

Calabria, Marco, Albert Costa, David W. Green & Jubin Abutalebi. 2018. Neural basis of bilingual language control. *Annals of the New York Academy of Sciences*.

Carl, Michael & Barbara Dragsted. 2017. Inside the monitor model: processes of default and challenged translation production. In Oliver Czulo & Silvia Hansen-Schirra (eds.), *Crossroads Between Contrastive Linguistics, Translation Studies And Machine Translation: Tc3 Ii*, 5–30: Zenodo.

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.

Chabasse, Catherine. 2009. *Gibt es eine Begabung für das Simultandolmetschen?*: *Erstellung eines Dolmetscheignungstests mit Schwerpunkt Simultandolmetschen*. Zugl.: Mainz, Univ. Diss. 2008 (SAXA. Beiträge zur Translationswissenschaft 4). Berlin: SAXA Verlag.

Chein, Jason M. & 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.

Chernov, Ghelly V. 1979. Semantic Aspects of Psycholinguistic research in Simultaneous Interpretation. *Language and Speech* 22(22 // 3). 277–295.

Chmiel, Agnieszka. 2021. Effects of simultaneous interpreting experience and training on anticipation, as measured by word-translation latencies. *Interpreting* 23(1). 18–44.

Christoffels, I. A. Degroot & J. 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.

Christoffels, Ingrid K. & Annette M.B. de Groot. 2005. Simultaneous Interpreting: A cognitive Perspective. In J.F. Kroll & A.M.B. de Groot (eds.), *Handbook of Bilingualism*, 454–479: Oxford University Press.

Coderre, Emily L. Jason F. Smith, Walter J. B. van Heuven & Barry Horwitz. 2016. The Functional Overlap of Executive Control and Language Processing in Bilinguals. *Bilingualism: Language and Cognition* 19(3). 471–488.

Coggins, Porter E. Teresa J. Kennedy & Terry A. Armstrong. 2004. Bilingual corpus callosum variability. *Brain and language* 89(1). 69–75.

Cohen, Jacob. 1977. *Statistical power analysis for the behavioral sciences*. New York, NY: Acad. Press.

Collette, Fabienne, Martial van der Linden, Steven Laureys, Guy Delfiore, Christian Degueldre, Andre Luxen & Eric Salmon. 2005. Exploring the unity and diversity of the neural substrates of executive functioning. *Human brain mapping* 25(4). 409–423.

Comishen, Kyle J. & Ellen Bialystok. 2021. Increases in attentional demands are associated with language group differences in working memory performance. *Brain and cognition* 147. 105658.

Compumedics Neuroscan. *SynAmps 2 system [EEG amplifier]* : Compumedics Neuroscan.

Conway, Andrew R. A. Michael J. Kane, Michael F. Bunting, D. Z. Hambrick, Oliver Wilhelm & Randall W. Engle. 2005. Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review* 12(5). 769–786.

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.

Costandi, Moheb. 2016. *Neuroplasticity* (The MIT Press essential knowledge series). Cambridge, Massachusetts: The MIT Press.

Cowan, Nelson. 1988. Evolving Conceptions of Memory Storage, Selective Attention, and Their Mutual Contrains Within the Human Information-Processing System. *Psychological bulletin*(104). 163–191.

Cowan, Nelson. 1998. *Attention and Memory* : Oxford University Press.

Cowan, Nelson. 2008. What are the differences between long-term, short-term, and working memory? In Wayne S. Sossin (ed.), *Essence of memory*, 1st edn. (Progress in Brain Research 169), 323–338. Amsterdam, Heidelberg: Elsevier.

Daffner, Kirk R. Hyemi Chong, Xue Sun, Elise C. Tarbi, Jenna L. Riis, Scott M. McGinnis & Phillip J. Holcomb. 2011. Mechanisms underlying age- and performance-related differences in working memory. *Journal of Cognitive Neuroscience* 23(6). 1298–1314.

de Bruin, Angela, Barbara Treccani & Sergio Della Sala. 2015. Cognitive advantage in bilingualism: an example of publication bias? *Psychological science* 26(1). 99–107.

de Groot, Annette M.B. 1997. The Cognitove Study of Translation and Interpreting. In Joseph H. Danks, Gregory M. Shreve, Stephen B. Fountain & Michael K. McBeath (eds.), *Cognitive processes in translation and interpreting*, 25–56. Thousand Oaks: Sage Publications.

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 J. Midgley, Phillip J. 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.

D'Esposito, Mark & Bradley R. Postle. 2015. The cognitive neuroscience of working memory. *Annual review of psychology* 66(Volume 66, 2015). 115–142.

Diamond, Adele. 2013. Executive functions. *Annual review of psychology* 64. 135–168.

Diamond, M. C. D. Krech & M. R. Rosenzweig. 1964. THE EFFECTS OF AN ENRICHED ENVIRONMENT ON THE HISTOLOGY OF THE RAT CEREBRAL CORTEX. *The Journal of comparative neurology* 123. 111–120.

Dijkstra, T. O. HENK van JAARSVELD & Sjoerd ten Brinke. 1998. Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition* 1(1). 51–66.

Dijkstra, T. O. Alexander Wahl, Franka Buytenhuijs, NINO van HALEM, Zina Al-Jibouri, Marcel de Korte & Steven Rekké. 2019. Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition* 22(04). 657–679.

Dijkstra, Ton & Walter J.B. van Heuven. 1998. The BIA model and bilingual word recognition. In J. Grainger & A.M. Jacobs (eds.), *Scientific psychology series. Localist connectionist approaches to human cognition*, 189–225: Lawrence Erlbaum Associates Publishers.

Dijkstra, Ton & Walter J.B. van Heuven. 2002. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5(3). 175–197.

Dong, Guangheng, Lizhu Yang, Yanbo Hu & Yue Jiang. 2009. Is N2 associated with successful suppression of behavior responses in impulse control processes? *Neuroreport* 20(6). 537–542.

Dong, Shanshan, Lynne M. Reder, Yuan Yao, Yuqiu Liu & Feiyan Chen. 2015. Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. *Brain research* 1616. 146–156.

Dong, Yanping & Yuhua Liu. 2016. Classes in Translating and Interpreting Produce Differential Gains in Switching and Updating. *Frontiers in psychology* 7. 1297.

Dong, Yanping & Zhilong Xie. 2014. Contributions of second language proficiency and interpreting experience to cognitive control differences among young adult bilinguals. *Journal of cognitive psychology (Hove, England)* 26(5). 506–519.

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.

Draganski, Bogdan, Christian Gaser, Volker Busch, Gerhard Schuierer, Ulrich Bogdahn & Arne May. 2004. Neuroplasticity: changes in grey matter induced by training. *Nature* 427(6972). 311–312.

Dunbar, K. & C. M. MacLeod. 1984. A horse race of a different color: Stroop interference patterns with transformed words. *Journal of experimental psychology. Human perception and performance* 10(5). 622–639.

EFFORT Project. 2023. Self-assessment tool. https://www.effortproject.eu/framework-for-translation-levels-a-b/self-assessment-tool/ (24 March, 2025).

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 & Jürg Kühnis. 2016. Functional Connectivity in the Left Dorsal Stream Facilitates Simultaneous Language Translation: An EEG Study. *Frontiers in human neuroscience* 10. 60.

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.

Emmorey, Karen, Gigi Luk, Jennie E. Pyers & Ellen Bialystok. 2008. The source of enhanced cognitive control in bilinguals: evidence from bimodal bilinguals. *Psychological science* 19(12). 1201–1206.

Ericsson, K. A. 2000. Expertise in interpreting. *Interpreting* 5(2). 187–220.

Eriksen, Barbara A. & Charles W. Eriksen. 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics* 16(1). 143–149.

Esposito, Alena G. Lynne Baker-Ward & Shane Mueller. 2013. Interference Suppression vs. Response Inhibition: An Explanation for the Absence of a Bilingual Advantage in Preschoolers' Stroop Task Performance. *Cognitive development* 28(4). 354–363.

European Commission. 2022. European Master's in Translation Competence Framework. https://commission.europa.eu/system/files/2022-11/emt_competence_fwk_2022_en.pdf (29 November, 2024).

Fabbro, F. & V. Darò. 1995. Delayed auditory feedback in polyglot simultaneous interpreters. *Brain and language* 48(3). 309–319.

Fabbro, Franco. 1999. *The Neurolinguistics of Bilingualism*: *An Introduction*. East Sussex: Psychology Press.

Fachbereich 06. 2021. Modulhandbuch des Studiengang M.A. (Master of Arts) Konferenzdolmetschen. https://studium.fb06.uni-mainz.de/files/2018/09/Modulhandbuch-MAKD.pdf (26 March, 2025).

Falkenstein, M. J. Hoormann & J. Hohnsbein. 1999. ERP components in Go/Nogo tasks and their relation to inhibition. *Acta psychologica* 101(2-3). 267–291.

Fan, Jin, Bruce D. McCandliss, Tobias Sommer, Amir Raz & Michael I. Posner. 2002. Testing the efficiency and independence of attentional networks. *Journal of cognitive neuroscience* 14(3). 340–347.

Felton, Adam, David Vazquez, Aurora I. Ramos-Nunez, Maya R. Greene, Alessandra McDowell, Arturo E. Hernandez & Christine Chiarello. 2017. Bilingualism Influences Structural Indices of Interhemispheric Organization. *Journal of Neurolinguistics* 42. 1–11.

Fillmore, Charles J. 1976. FRAME SEMANTICS AND THE NATURE OF LANGUAGE*. *Annals of the New York Academy of Sciences* 280(1). 20–32.

Finkbeiner, Matthew, Jorge Almeida, Niels Janssen & Alfonso Caramazza. 2006. Lexical selection in bilingual speech production does not involve language suppression. *Journal of experimental psychology. Learning, memory, and cognition* 32(5). 1075–1089.

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.

Frauenfelder, Ulrich & Herbert Schriefers. 1997. A psycholinguistic perspective on Simultaneous Interpretation. *Interpreting* 2(1-2). 55–89.

Garbin, G. A. Sanjuan, C. Forn, J. C. Bustamante, A. Rodriguez-Pujadas, V. Belloch, M. Hernandez, A. Costa & C. Avila. 2010. Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. *NeuroImage* 53(4). 1272–1278.

Garcia, Adolfo M. 2019. *The neurocognition of translation and interpreting* (Benjamins translation library volume 147). Amsterdam, Philadelphia: John Benjamins Publishing Company.

García, Adolfo M. 2014. The interpreter advantage hypothesis. *Translation and Interpreting Studies* 9(2). 219–238.

García, Adolfo M. 2015. Psycholinguistic explorations of lexical translation equivalents: Thirty years of research and their implications for cognitive translatology. *Translation Spaces* 4(1). 9–28.

García, Adolfo M. Edinson Muñoz & Boris Kogan. 2020. Taxing the bilingual mind: Effects of simultaneous interpreting experience on verbal and executive mechanisms. *Bilingualism: Language and Cognition* 23(4). 729–739.

García-Pentón, Lorna, Yuriem Fernández García, Brendan Costello, Jon A. Duñabeitia & Manuel Carreiras. 2016. The neuroanatomy of bilingualism: how to turn a hazy view into the full picture. *Language, Cognition and Neuroscience* 31(3). 303–327.

Gerver, David. 1976. Empirical Studies of Simultaneous Interpretation: A Review and a Model. In Richard W. Brislin (ed.), *Translation - Applications and Research*: *Applications and research*, 165–207. New York: Gardner Press distributed by Halsted Press.

Gerver, David. 2002. The effects of source language presentation rate on the performance of simultaneous conference interpreters. In Franz Pöchhacker & Miriam Shlesinger (eds.), *The interpreting studies reader*, 1st edn. (Routledge language readers), 52-66. London: Routledge.

Gieshoff, Anne Catherine. 2017. Audiovisual speech decreases the number of cognate translations in simultaneous interpreting. In Silvia Hansen-Schirra, Oliver Czulo & Sascha Hofmann (eds.), *Empirical modelling of translation and interpreting* (Translation and Multilingual Natural Language Processing 7), 313–330. Berlin, Germany: Language Science Press.

Gile, Daniel. 2009. *Basic concepts and models for interpreter and translator training* (Benjamins translation library EST suberies v. 8). Amsterdam, Philadelphia: John Benjamins Pub. Co.

Gile, Daniel. 2015. The contributions of cognitive psychology and psycholinguistics to conference interpreting: A critical anaylsis. In Aline Ferreira & John W. Schwieter (eds.), *Psycholinguistic and Cognitive Inquiries into Translation and Interpreting* (EST Subseries 15), 41–64. Amsterdam, New York: John Benjamins Publishing Company.

Gold, Brian T. Chobok Kim, Nathan F. Johnson, Richard J. Kryscio & Charles D. Smith. 2013. Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *Journal of Neuroscience* 33(2). 387–396.

Göpferich, Susanne. 2009. Towards a model of translation competence and its acquisition: the longitudinal study TransComp. In Arnt Lykke Jakobsen, Inger M. Mees & Susanne Göpferich (eds.), *Behind the Mind : Copenhagen Studies in Language* (37), 11–37. Frederiksberg: Samfundslittertur.

Grady, Cheryl L. Gigi Luk, Fergus I. M. Craik & Ellen Bialystok. 2015. Brain network activity in monolingual and bilingual older adults. *Neuropsychologia* 66. 170–181.

Green, C. S. & Daphne Bavelier. 2003. Action video game modifies visual selective attention. *Nature* 423(6939). 534–537.

Green, David W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1(2). 67–81.

Green, David W. & Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of cognitive psychology (Hove, England)* 25(5). 515–530.

Green, David W. & Li Wei. 2014. A control process model of code-switching. *Language, Cognition and Neuroscience* 29(4). 499–511.

Grogan, A. O. Parker Jones, N. Ali, J. Crinion, S. Orabona, M. L. Mechias, S. Ramsden, D. W. Green & C. J. Price. 2012. Structural correlates for lexical efficiency and number of languages in non-native speakers of English. *Neuropsychologia* 50(7). 1347–1352.

Grosjean, F. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and language* 36(1). 3–15.

Grosjean, François & Ping Li. 2013. *The Psycholinguistics of Bilingualism* : Wiley Blackwell Publishing.

Grundy, John G. John A. E. Anderson & Ellen Bialystok. 2017a. Neural correlates of cognitive processing in monolinguals and bilinguals. *Annals of the New York Academy of Sciences* 1396(1). 183–201.

Grundy, John G. Ashley Chung-Fat-Yim, Deanna C. Friesen, Lorinda Mak & Ellen Bialystok. 2017b. Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition* 163. 42–55.

Hallett, P. E. 1978. Primary and secondary saccades to goals defined by instructions. *Vision research* 18(10). 1279–1296.

Halverson, Sandra L. 2003. The cognitive basis of translation universals. *Target. International Journal of Translation Studies* 15(2). 197–241.

Halverson, Sandra L. 2017. Gravitational pull in translation. Testing a revised model. In Geert de Sutter, Marie-Aude Lefer & Isabelle Delaere (eds.), *Empirical Translation Studies : New Methodological and Theoretical Traditions*, 9–45. Berlin: De Gruyter Mouton.

Hansen, Laura B. Pedro Macizo, Jon A. Duñabeitia, David Saldaña, Manuel Carreiras, Luis J. Fuentes & M. T. Bajo. 2016. Emergent Bilingualism and Working Memory Development in School Aged Children. *Language Learning* 66(S2). 51–75.

Hebb, Donald O. 1949. *The Organization of Behaviour*: *A Neuropsychological Theory*. New York, London: John Wiley & Sons, Inc.

Heil, Martin, Allen Osman, Juliane Wiegelmann, Bettina Rolke & Erwin Hennighausen. 2000. N200 in the Eriksen-Task: Inhibitory Executive Processes? *Journal of Psychophysiology* 14(4). 218–225.

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, A. E. A. Martinez & K. 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.

Hernández, Mireia, Clara D. Martin, Francisco Barceló & Albert Costa. 2013. Where is the bilingual advantage in task-switching? *Journal of Memory and Language* 69(3). 257–276.

Hervais-Adelman, Alexis & Laura Babcock. 2020. The neurobiology of simultaneous interpreting: Where extreme language control and cognitive control intersect. *Bilingualism: Language and Cognition* 23(4). 740–751.

Hervais-Adelman, Alexis, Barbara Moser-Mercer & Narly Golestani. 2015a. Brain functional plasticity associated with the emergence of expertise in extreme language control. *NeuroImage* 114. 264–274.

Hervais-Adelman, Alexis, Barbara Moser-Mercer, Christoph M. Michel & Narly Golestani. 2015b. fMRI of Simultaneous Interpretation Reveals the Neural Basis of Extreme Language Control. *Cerebral cortex (New York, N.Y. : 1991)* 25(12). 4727–4739.

Hervais-Adelman, Alexis, Barbara Moser-Mercer, Micah M. Murray & Narly Golestani. 2017. Cortical thickness increases after simultaneous interpretation training. *Neuropsychologia* 98. 212–219.

Hervais-Adelman, Alexis G. Barbara Moser-Mercer & Narly Golestani. 2011. Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. *Frontiers in psychology* 2. 234.

Hilchey, Matthew D. & Raymond M. Klein. 2011. Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic bulletin & review* 18(4). 625–658.

Ho, Yim-Chi, Mei-Chun Cheung & Agnes S. Chan. 2003. Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology* 17(3). 439–450.

Hoffman, Robert R. 1997. The cognitive psychology of expertise and the domain of interpreting. *Interpreting* 2(1-2). 189–230.

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.

Hu, Minxia & Wenjiao Fan. 2021. The interpreter advantage in executive functions—A systematic review and meta-analysis. *Forum for Linguistic Studies (Transferred)* 3(1). 131.

Jacob, Gunnar, Moritz J. 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* 45(3). 542–566.

Jaeggi, Susanne M. Martin Buschkuehl, Walter J. Perrig & Beat Meier. 2010. The concurrent validity of the N-back task as a working memory measure. *MEMORY* 18(4). 394–412.

Jakobson, Roman. 1959. On linguistic aspects of translation. In Reuben A. Brower (ed.), *On Translation*, 232–239: Harvard University Press.

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.

Jersild, Arthur T. 1927. *Mental Set And Shift* (Archives of Psychology 89). New York.

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.

Johnstone, Stuart J. Robert J. Barry, Valentina Markovska, Aneta Dimoska & Adam R. Clarke. 2009. Response inhibition and interference control in children with AD/HD: a visual ERP investigation. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* 72(2). 145–153.

Just, Marcel A. & Patricia A. Carpenter. 1980. A theory of reading: From eye fixations to comprehension. *Psychological review* 87(4). 329–354.

Kahneman, Daniel. 1973. *Attention and effort* (Prentice-Hall series in experimental psychology). Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Kalina, Sylvia. 2000. Interpreting competences as a basis and a goal for teaching. *The Interpreters' Newsletter*(10). 3–32.

Kalina, Sylvia. 2006. Zur Dokumentation von Maßnahmen der Qualitätssicherung beim Konferenzdolmetschen. In Carmen Heine, Klaus Schubert & Heidrun Gerzymisch (eds.), *Text and translation*: *Theory and methodology of translation : [Beiträge der Saarbrücker Euroconference 2004* (Jahrbuch Übersetzen und Dolmetschen 6), 253–268. Tübingen: Narr.

Karayanidis, Frini & Sharna D. Jamadar. 2014. Event-Related Potentials Reveal Multiple Components of Proactive and Reactive Control in Task Switching. In James Grange & George Houghton (eds.), *Task switching and cognitive control*, 200–236. Oxford: Oxford University Press.

Khateb, Asaid, Jubin Abutalebi, Christoph M. Michel, Alan J. Pegna, Hannelore Lee-Jahnke & Jean-Marie Annoni. 2007. Language selection in bilinguals: a spatio-temporal analysis of electric brain activity. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* 65(3). 201–213.

Kimball, D.R. & K.L. Holyoak. 2000. Transfer and expertise. In Tulving E. & Craik F.I.M. (eds.), *The Oxford handbook of memory*, 109–122: Oxford University Press.

Kleinsorge, Thomas. 2021. Stimulus-Response Conflict Tasks and Their Use in Clinical Psychology. *International journal of environmental research and public health* 18(20).

Klimesch, Wolfgang, Paul Sauseng & Simon Hanslmayr. 2007. EEG alpha oscillations: the inhibition-timing hypothesis. *Brain research reviews* 53(1). 63–88.

Koechlin, Etienne, Chrystèle Ody & Frédérique Kouneiher. 2003. The architecture of cognitive control in the human prefrontal cortex. *Science (New York, N.Y.)* 302(5648). 1181–1185.

Kok, A. 2001. On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* 38(3). 557–577.

Konorski, Jerzy. 1948. *Conditioned reflexes and neuron organization*. Cambridge: Cambridge University Press.

Köpke, Barbara & Jean-Luc Nespoulous. 2006. Working memory performance in expert and novice interpreters. *Interpreting. International Journal of Research and Practice in Interpreting* 8(1). 1–23.

Köpke, Barbara & Teresa M. Signorelli. 2012. Methodological aspects of working memory assessment in simultaneous interpreters. *International Journal of Bilingualism* 16(2). 183–197.

Kopp, B. F. Rist & U. Mattler. 1996. N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology* 33(3). 282–294.

Korenar, M. J. Treffers-Daller & C. Pliatsikas. 2023. Brain structure adapts dynamically to highly demanding bilingual experiences: Insights from interpreters and translators. *Ampersand* 11.

Kousaie, Shanna, Xiaoqian J. Chai, Kaija M. Sander & Denise Klein. 2017. Simultaneous learning of two languages from birth positively impacts intrinsic functional connectivity and cognitive control. *Brain and cognition* 117. 49–56.

Kousaie, Shanna & Natalie A. Phillips. 2012. Conflict monitoring and resolution: are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain research* 1446.

Kousaie, Shanna & Natalie A. Phillips. 2017. A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia* 94. 23–35.

Krampe, R. T. & K. A. Ericsson. 1996. Maintaining excellence: deliberate practice and elite performance in young and older pianists. *Journal of experimental psychology. General* 125(4). 331–359.

Kroll, J. F. & E. Stewart. 1994. Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language* 33(2). 149–174.

Kutz, Wladimir. 2010. *Dolmetschkompetenz*: *Was muss der Dolmetscher wissen und können?*. Berlin, London, Bochum, Paris: European University Press.

Kuzyk, Olivia, Margaret Friend, Vivianne Severdija, Pascal Zesiger & Diane Poulin-Dubois. 2020. Are there Cognitive Benefits of Code-switching in Bilingual Children? A longitudinal study. *Bilingualism: Language and Cognition* 23(3). 542–553.

Lacoste, M. C. de, J. B. Kirkpatrick & E. D. Ross. 1985. Topography of the human corpus callosum. *Journal of neuropathology and experimental neurology* 44(6). 578–591.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar 1. Theoretical prerequisites*. Stanford, Calif.: Stanford Univ. Press.

Lavric, Aureliu, Diego A. Pizzagalli & Simon Forstmeier. 2004. When 'go' and 'nogo' are equally frequent: ERP components and cortical tomography. *The European journal of neuroscience* 20(9). 2483–2488.

Lee, Tae-Hyung. 2002. Ear Voice Span in English into Korean Simultaneous Interpretation. *Meta* 47(4). 596–606.

Lehtonen, Minna H. Matti Laine, Jussi Niemi, Tormod Thomsen, Victor A. Vorobyev & Kenneth Hugdahl. 2005. Brain correlates of sentence translation in Finnish-Norwegian bilinguals. *Neuroreport* 16(6). 607–610.

Levelt, Willem J. M. 1995. The ability to speak: from intentions to spoken words. *European Review* 3(1). 13–23.

Levelt, Willem J. M. 1998. *Speaking*: *From intention to articulation* (ACL-MIT Press series in natural-language processing), 5th edn. Cambridge, Mass.: MIT Press.

Levelt, Willem J.M. 1989. *Speaking*: *From Intention to Articulation*. Cambridge, Mass.: MIT Press.

Levelt, Willem J.M. 1999. Producing spoken language: a blueprint of the speaker. In Colin M. Brown & Peter Hagoort (eds.), *The Neurocognition of Language*, 83–122. New York: Oxford University Press.

Lewis-Peacock, Jarrod A. Andrew T. Drysdale, Klaus Oberauer & Bradley R. Postle. 2012. Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience* 24(1). 61–79.

López Gómez, María J. Teresa Bajo Molina, Presentación Padilla Benítez & Julio Santiago de Torres. 2007. *Predicting proficiency in signed language interpreting* (Interpreting 9): John Benjamins Publishing Company.

López Zunini, Rocío A. Cassandra Morrison, Shanna Kousaie & Vanessa Taler. 2019. Task switching and bilingualism in young and older adults: A behavioral and electrophysiological investigation. *Neuropsychologia* 133.

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.

Luck, Steven J. 2005. *An introduction to the event-related potential technique* (Cognitive neuroscience). Cambridge, Mass.: MIT Press.

Luck, Steven J. & Emily S. Kappenmann. 2012. *The Oxford handbook of event-related potential components* (Oxford library of psychology). New York, Oxford: Oxford University Press.

Luk, Gigi, John A. E. Anderson, Fergus I. M. Craik, Cheryl Grady & Ellen Bialystok. 2010. Distinct neural correlates for two types of inhibition in bilinguals: response inhibition versus interference suppression. *Brain and cognition* 74(3). 347–357.

Luk, Gigi, Ellen Bialystok, Fergus I. M. Craik & Cheryl L. Grady. 2011. Lifelong bilingualism maintains white matter integrity in older adults. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31(46). 16808–16813.

Macnamara, Brooke N. & Andrew R. A. Conway. 2014. Novel evidence in support of the bilingual advantage: influences of task demands and experience on cognitive control and working memory. *Psychonomic Bulletin & Review* 21(2). 520–525.

Maguire, E. A. D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. Frackowiak & C. D. Frith. 2000. Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America* 97(8). 4398–4403.

Makeig, S. 1993. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and Clinical Neurophysiology* 86(4). 283–293.

Makeig, Scott, Stefan Debener, Julie Onton & Arnaud Delorme. 2004. Mining event-related brain dynamics. *Trends in cognitive sciences* 8(5). 204–210.

Marian, Viorica, Henrike K. Blumenfeld & Margarita Kaushanskaya. 2007. The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of speech, language, and hearing research : JSLHR* 50(4). 940–967.

Marian, Viorica & Sayuri Hayakawa. 2021. Measuring Bilingualism: The Quest for a "Bilingualism Quotient". *Applied Psycholinguistics* 42(Suppl 2). 527–548.

Markiewicz, Roksana, Ali Mazaheri & Andrea Krott. 2021. Bilingual speakers’ enhanced monitoring can slow them down 87.

Markiewicz, Roksana, Ali Mazaheri & Andrea Krott. 2023. Bilingualism can cause enhanced monitoring and occasional delayed responses in a flanker task. *The European journal of neuroscience* 57(1). 129–147.

Martin-Rhee, Michelle M. & Ellen Bialystok. 2008. The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition* 11(1). 81–93.

Martin-Rhee, Michelle M. & Ellen Bialystrok. 2008. The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition* 11(1). 81–93.

Mechelli, Andrea, Jenny T. 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.

Meuter, Renata F. & Alan Allport. 1999. Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language* 40(1). 25–40.

Miller, E. K. & J. D. Cohen. 2001. An integrative theory of prefrontal cortex function. *Annual review of neuroscience* 24. 167–202.

Miyake, A. N. P. Friedman, M. J. Emerson, A. H. Witzki, A. Howerter & T. D. 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.

Miyake, Akira & Naomi P. Friedman. 2012. The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current directions in psychological science* 21(1). 8–14.

Mohades, Seyede G. Esli Struys, Peter van Schuerbeek, Chris Baeken, Piet van de Craen & Robert Luypaert. 2014. Age of second language acquisition affects nonverbal conflict processing in children: an fMRI study. *Brain and behavior* 4(5). 626–642.

Monsell, Stephen. 2003. Task switching. *Trends in Cognitive Sciences* 7(3). 134–140.

Morales, Julia, Francisca Padilla, Carlos J. Gómez-Ariza & M. T. Bajo. 2015. Simultaneous interpretation selectively influences working memory and attentional networks. *Acta psychologica* 155. 82–91.

Moreno, Sylvain, Ellen Bialystok, Raluca Barac, E. G. Schellenberg, Nicholas J. Cepeda & Tom Chau. 2011. Short-term music training enhances verbal intelligence and executive function. *Psychological science* 22(11). 1425–1433.

Morrison, Cassandra, Farooq Kamal & Vanessa Taler. 2019. The influence of bilingualism on working memory event-related potentials. *Bilingualism: Language and Cognition* 22(1). 191–199.

Morrison, Cassandra & Vanessa Taler. 2020. ERP measures of the effects of age and bilingualism on working memory performance. *Neuropsychologia* 143. 107468.

Moser, Barbara. 1978. Simultaneous Interpretation: A Hypothetical Model and its Practical Application. In David Gerver & H. Wallace Sinaiko (eds.), *Language Interpretation and Communication* (NATO conference series v. 6), 353–368. New York, London: Plenum Press.

Moser-Mercer, Barbara. 2008. Skill Acquisition in Interpreting. *The Interpreter and Translator Trainer* 2(1). 1–28.

Moser-Mercer, Barbara, Ulrich Hans Frauenfelder, Beatriz Casado & Alexander Künzli. 2000. Searching to define expertise in interpreting. In B. Englund Dimitrova & K. Hyltenstam (eds.), *Language processing and simultaneous interpreting: Interdisciplinary perspectives*, 107–131: John Benjamins Publishing Company.

Muñoz Martín, Ricardo. 2022. Cognitive translation and interpreting studies. https://www.aieti.eu/enti/ctis_ENG/entry.html (16 December, 2024).

Muysken, Pieter. 2000. *Bilingual Speech*: *A Typology of Code-Mixing*. Cambridge: Cambridge University Press.

Nichols, Emily S. & Marc F. Joanisse. 2016. Functional activity and white matter microstructure reveal the independent effects of age of acquisition and proficiency on second-language learning. *NeuroImage* 143. 15–25.

Nieuwenhuis, Sander, Nick Yeung, Wery van den Wildenberg & K. R. Ridderinkhof. 2003. Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, affective & behavioral neuroscience* 3(1). 17–26.

Norman, Donald A. & Tim Shallice. 1986. Attention to Action. In Richard J. Davidson, Gary E. Schwartz & David Shapiro (eds.), *Consciousness and Self-Regulation*: *Advances in Research and Theory Volume 4*, 1–18. Boston, MA, s.l.: Springer US.

Nour, Soudabeh, Esli Struys, Evy Woumans, Ily Hollebeke & Hélène Stengers. 2020. An interpreter advantage in executive functions? *Interpreting. International Journal of Research and Practice in Interpreting* 22(2). 163–186.

Olulade, O. A. N. I. Jamal, D. S. Koo, C. A. Perfetti, C. LaSasso & G. F. Eden. 2016. Neuroanatomical Evidence in Support of the Bilingual Advantage Theory. *Cerebral cortex (New York, N.Y. : 1991)* 26(7). 3196–3204.

Paap, Kenneth R. 2018. Bilingualism in Cognitive Science. In Annick de Houwer & Lourdes Ortega (eds.), *The Cambridge Handbook of Bilingualism*, 435–465: Cambridge University Press.

Paap, Kenneth R. & Zachary I. Greenberg. 2013. There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive psychology* 66(2). 232–258.

Paap, Kenneth R. Hunter A. Johnson & Oliver Sawi. 2015. Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex; a journal devoted to the study of the nervous system and behavior* 69. 265–278.

Paap, Kenneth R. Hunter A. Myuz, Regina T. Anders, Morgan F. Bockelman, Roman Mikulinsky & Oliver M. Sawi. 2017. No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of cognitive psychology (Hove, England)* 29(2). 89–112.

PACTE. 2000. Acquiring Translation Competence: Hypotheses and Methodological Problems of a Research Project. In Allison Beeby, Doris Ensinger & Marisa Presas (eds.), *Investigating Translation*: *Selectec papers from the 4th international congress on translation, Barcelona, 1998*, 99–106. Amsterdam/Philadephia: John Benjamins Publishing Company.

PACTE group, A. Beeby, M.F. Rodriguez, O. Fox, A.H. Albir, W. Neunzig, M. Orozco, M. Presas, P.R. Ines & L. Romero. 2003. Building a translation competence model: Triangulating Translation Perspectives in process oriented research. In Fabio Alves (ed.), *Triangulating Translation* (45), 43–68. Amsterdam: John Benjamins Publishing Company.

Paradis, M. M. C. Goldblum & R. Abidi. 1982. Alternate antagonism with paradoxical translation behavior in two bilingual aphasic patients. *Brain and language* 15(1). 55–69.

Paradis, Michel. 1994. Toward a neurolinguistic theory of simultaneous translation: The framework. *nternational Journal of Psycholinguistics*(10(3)). 319–335.

Peal, Elizabeth & Wallace E. Lambert. 1962. The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied* 76(27). 1–23.

Periáñez, Jose A. & Francisco Barceló. 2009. Updating sensory versus task representations during task-switching: insights from cognitive brain potentials in humans. *Neuropsychologia* 47(4). 1160–1172.

Pfefferbaum, Adolf, Judith M. Ford, Barbara J. Weller & Bert S. Kopell. 1985. ERPs to response production and inhibition. *Electroencephalography and Clinical Neurophysiology* 60(5). 423–434.

Pfurtscheller, G. 1977. Graphical display and statistical evaluation of event-related desynchronization (ERD). *Electroencephalography and Clinical Neurophysiology* 43(5). 757–760.

Pfurtscheller, G. & F. H. Da Lopes Silva. 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology* 110(11). 1842–1857.

Pliatsikas, Christos. 2019. Multilingualism and Brain Plasticity. In John W. Schwieter & Michel Paradis (eds.), *The handbook of the neuroscience of multilingualism* (Blackwell Handbooks in Linguistics), 230–251. Hoboken, NJ: Wiley Blackwell.

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 D. Saddy. 2017. Immersive bilingualism reshapes the core of the brain. *Brain structure & function* 222(4). 1785–1795.

Pöchhacker, Franz. 2016. *Introducing interpreting studies*. London, New York: Routledge.

Polich, John. 2007. Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology* 118(10). 2128–2148.

Posner, Michel I. & Gregory J. DiGirolamo. 1998. Executive attention: Conflict, target detection, and cognitive control. In Raja Parasuraman (ed.), *The attentive brain*, 401–423: MIT Press.

Postle, Bradley R. T. J. Druzgal & Mark D'Esposito. 2003. Seeking the neural substrates of visual working memory storage. *Cortex; a journal devoted to the study of the nervous system and behavior* 39(4-5). 927–946.

Price, C. J. D. W. Green & R. 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 H. Gollan. 2011. Good language-switchers are good task-switchers: evidence from Spanish-English and Mandarin-English bilinguals. *Journal of the International Neuropsychological Society : JINS* 17(4). 682–691.

Prior, Anat & Brian Macwhinney. 2010. A bilingual advantage in task switching. *Bilingualism: Language and Cognition* 13(2). 253–262.

Privitera, Adam J. Mohammad Momenian & Brendan S. Weekes. 2023. Modeling the bilingual advantage: Do results differ between analysis methods? *Ampersand* 11. 100134.

Purmann, Sascha, Stephanie Badde, Aquiles Luna-Rodriguez & Mike Wendt. 2011. Adaptation to Frequent Conflict in the Eriksen Flanker Task. *Journal of Psychophysiology* 25(2). 50–59.

Purves, Dale, George J Augustine, David Fitzpatrick, William C Hall, Anthony-Samuel LaMantia & Leonard E. White. 2012. *Neuroscience*, 5th edn. Sunderland, Mass.: Sinauer.

Quallo, M. M. C. J. Price, K. Ueno, T. Asamizuya, K. Cheng, R. N. Lemon & A. Iriki. 2009. Gray and white matter changes associated with tool-use learning in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America* 106(43). 18379–18384.

Quaresima, Valentina, Marco Ferrari, Marco C. P. van der Sluijs, Jan Menssen & Willy N. J. M. Colier. 2002. Lateral frontal cortex oxygenation changes during translation and language switching revealed by non-invasive near-infrared multi-point measurements. *Brain research bulletin* 59(3). 235–243.

Riccardi, Alessandra. 2005. On the Evolution of Interpreting Strategies in Simultaneous Interpreting. *Meta* 50(2). 753–767.

Riehl, Claudia M. 2015. *Mehrsprachigkeit*: *Eine Einführung*. Darmstadt: WBG - Wissenschaftliche Buchgesellschaft.

Rinne, J. O. J. Tommola, M. Laine, B. J. Krause, D. Schmidt, V. Kaasinen, M. Teräs, H. Sipilä & M. Sunnari. 2000. The translating brain: cerebral activation patterns during simultaneous interpreting. *Neuroscience letters* 294(2). 85–88.

Rodriguez-Fornells, Antoni, Ulrike M. Krämer, Urbano Lorenzo-Seva, Julia Festman & Thomas F. Münte. 2011. Self-assessment of individual differences in language switching. *Frontiers in Psychology* 2. 388.

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 F. Kroll, Michele T. Diaz & Sharlene D. Newman. 2017. Changes in White-Matter Connectivity in Late Second Language Learners: Evidence from Diffusion Tensor Imaging. *Frontiers in Psychology* 8(2040). 1–15.

Rushworth, Matthew F. S. R. E. Passingham & A. C. Nobre. 2002. Components of switching intentional set. *Journal of Cognitive Neuroscience* 14(8). 1139–1150.

Schaeffer, Moritz & Michael Carl. 2015. Shared representations and the translation process: A recursive model. In Maureen Ehrensberger-Dow, Birgitta Englund Dimitrova, Séverine Hubscher-Davidson & Ulf Norberg (eds.), *Describing Cognitive Processes in Translation* (John Benjamins Current Topics 77), 21–42. Amsterdam: John Benjamins.

Schaeffer, Moritz, David Huepe, Silvia Hansen-Schirra, Sascha Hofmann, Edinson Muñoz, Boris Kogan, Eduar Herrera, Agustín Ibáñez & Adolfo M. García. 2020. The Translation and Interpreting Competence Questionnaire: an online tool for research on translators and interpreters. *Perspectives* 28(1). 90–108.

Schellenberg, E. G. 2004. Music lessons enhance IQ. *Psychological science* 15(8). 511–514.

Schwieter, John W. & Julia Festman. 2023. *The cognitive neuroscience of bilingualism* (Cambridge fundamentals of neuroscience in psychology). Cambridge, New York, NY: Cambridge University Press.

Sebastián-Gallés, Núria, Bàrbara Albareda-Castellot, Whitney M. Weikum & Janet F. Werker. 2012. A bilingual advantage in visual language discrimination in infancy. *Psychological science* 23(9). 994–999.

Seeber, Kilian G. 2007. Thinking outside the cube: modeling language processing tasks in a multiple resource paradigm. *Proc. Interspeech*. 1382–1385.

Seeber, Kilian G. 2011. Cognitive load in simultaneous interpreting. *Interpreting* 13(2). 176–204.

Seeber, Kilian G. 2017. 25 Multimodal Processing in Simultaneous Interpreting. In John W. Schwieter & Aline Ferreira (eds.), *The Handbook of Translation and Cognition*, 461–475: Wiley.

Seleskovitch, Danica & Marianne Lederer. 1989. *Pédagogie raisonnée de l'interprétation*. Bruxelles-Luxembourg: Didier Érudition.

Selya, Arielle S. Jennifer S. Rose, Lisa C. Dierker, Donald Hedeker & Robin J. 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.

Setton, Robin. 1999. *Simultaneous Interpretation*: *a cognitive-pragmatic analysis* (Benjamins translation library). Amsterdam: John Benjamins Publishing Company.

Shallice, T. & E. K. Warrington. 1970. Independent functioning of verbal memory stores: a neuropsychological study. *The Quarterly journal of experimental psychology* 22(2). 261–273.

Shiffrin, R. M. & R. C. Atkinson. 1969. Storage and retrieval processes in long-term memory. *Psychological review* 76(2). 179–193.

Simon, J. R. & James D. Wolf. 1963. Choice reaction times as a function of angular stimulus-response correspondence and age. *Ergonomics* 6(1). 99–105.

Smith, E. E. & J. Jonides. 1997. Working memory: a view from neuroimaging. *Cognitive psychology* 33(1). 5–42.

Sternberg, Robert J. 2005. Intelligence, Competence, and Expertise. In Andrew J. Elliot & Carol S. Dweck (eds.), *Handbook of Competence and Motivation*, 15–30. New York, London: The Guilford Press.

Sulpizio, Simone, Nicola Del Maschio, Gianpaolo Del Mauro, Davide Fedeli & Jubin Abutalebi. 2020. Bilingualism as a gradient measure modulates functional connectivity of language and control networks. *NeuroImage* 205. 116306.

Sutton, S. M. Braren, J. Zubin & E. R. John. 1965. Evoked-potential correlates of stimulus uncertainty. *Science (New York, N.Y.)* 150(3700). 1187–1188.

Szmalec, Arnaud, Frederick Verbruggen, André Vandierendonck & Eva Kemps. 2011. Control of interference during working memory updating. *Journal of experimental psychology. Human perception and performance* 37(1). 137–151.

Tao, Lily, Gongting Wang, Miaomiao Zhu & Qing Cai. 2021. Bilingualism and domain-general cognitive functions from a neural perspective: A systematic review. *Neuroscience and biobehavioral reviews* 125. 264–295.

The MathWorks Inc. 2022. *MATLAB*. Natik, Massachusetts.

Timarová, Šárka, Barbara Dragsted & Inge Gorm Hansen. 2011. Time lag in translation and interpreting: A methodological exploration. In Cecilia Alvstad, Adelina Hild & Elisabeth Tiselius (eds.), *Methods and Strategies of Process Research*: *Integrative approaches in Translation Studies*, 221-146. Amsterdam/Philadephia: John Benjamins Publishing Company.

Timmer, Kalinka, John G. Grundy & Ellen Bialystok. 2017. Earlier and more distributed neural networks for bilinguals than monolinguals during switching. *Neuropsychologia* 106. 245–260.

Tirkkonen-Condit, Sonja. 2005. The Monitor Model Revisited: Evidence from Process Research. *Meta* 50(2). 405–414.

Tiselius, Elisabet & Adelina Hild. 2017. Expertise and Competence in Translation and Interpreting. In, *The Handbook of Translation and Cognition*, 423–444: John Wiley & Sons, Ltd.

Toury, Gideon. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam/Philadephia: John Benjamins.

Treisman, Anne M. 1965. The effects of redundancy and familiarity on translating and repeating back a foreign and a native language. *British Journal of Psychology*(56). 369–379.

Turner, Robert. 2019. Finding likeness: Neural plasticity and ritual experience. *Anthropology Today* 35(3). 3–6.

Tzou, Yeh-Zu, Zohreh R. Eslami, Hsin-Chin Chen & Jyotsna Vaid. 2012. Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: Evidence from Mandarin–English speakers. *International Journal of Bilingualism* 16(2). 213–227.

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 (Basel, Switzerland)* 9(3).

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 Vandelanotte, Arnaud Szmalec & Wouter Duyk. 2016. The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition* 19(1). 181–190.

Vilberg, Kaia L. & Michael D. Rugg. 2008. Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia* 46(7). 1787–1799.

Walter, W. G. R. Cooper, V. J. Aldridge, W. C. Mccallum & A. L. Winter. 1964. CONTINGENT NEGATIVE VARIATION: AN ELECTRIC SIGN OF SENSORIMOTOR ASSOCIATION AND EXPECTANCY IN THE HUMAN BRAIN. *Nature* 203(203). 380–384.

Wild-Wall, Nele, Michael Falkenstein & Joachim Hohnsbein. 2008. Flanker interference in young and older participants as reflected in event-related potentials. *Brain research* 1211. 72–84.

Winter, Bodo. 2020. *Statistics for Linguists: An Introduction Using R*. New York, NY: Routledhe, Taylor & Francis.

Woumans, Evy, Evy Ceuleers, Lize van der Linden, Arnaud Szmalec & Wouter Duyck. 2015. Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of experimental psychology. Learning, memory, and cognition* 41(5). 1579–1586.

Woumans, Evy & Wouter Duyck. 2015. The bilingual advantage debate: Moving toward different methods for verifying its existence. *Cortex; a journal devoted to the study of the nervous system and behavior* 73. 356–357.

Yu, Zhibin & Yanping Dong. 2022. The Emergence of a Complex Language Skill: Evidence from the Self-organization of Interpreting Competence in Interpreting Students. *Bilingualism: Language and Cognition* 25(2). 269–282.

Yudes, Carolina, Pedro Macizo & Teresa Bajo. 2011. The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in psychology* 2. 309.

Zhao, Hongming, Xiaocong Chen & Yanping Dong. 2024. Young interpreting trainees' better adaptation to the flanker conflicting environment: An ERP study. *Journal of Neurolinguistics* 70. 101181.

Zhong, Fei & Yanping Dong. 2024. ERP neural correlates of the interpreter advantage in coordination for interpreting students. *Applied Psycholinguistics* 45(5). 897–919.

Zuur, Alain F. Elena N. Ieno & Chris S. Elphick. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1(1). 3–14.

Appendix

Baseline TICQ Experimental group

Please enter the participant ID. Your participant ID is the number you have been given by the experimenter (example: P03).	Please enter the name of your study: Please use the study name EEGHa2021.	Module A1: Personal data Age:	Gender:	What is your highest degree?	What is your highest degree? [Other]	Which languages do you speak? [Mother tongue (L1)]	Which languages do you speak? [Language 2 (L2)]
EEGHa2021	P01	20	Female	Bachelor		Deutsch	Englisch
EEGHa2021	P02	25	Female	Bachelor		Deutsch	Englisch
EEGHa2021	P03	23	Female	Bachelor		Deutsch	Englisch
EEGHa2021	P04	29	Female	Master		Deutsch	Französisch
EEGHa2021	P06	23	Female	Bachelor		Deutsch	Englisch
EEGHa2021	P07	26	Female	Bachelor		Deutsch	Englisch
EEGHa2021	P08	22	Female	Bachelor		Deutsch	Englisch
EEGHa2021	P10	24	Female	Bachelor		Deutsch	Englisch
EEGHa2021	P11	24	Female	Other	Derzeitig noch am Schreiben der Bachelorarbeit	Deutsch	Englisch

Which languages do you speak? [Language 3 (L3)]	Which languages do you speak? [Language 4 (L4)]	Which languages do you speak? [Language 5 (L5)]	In which of these languages were you fluent before age 7? (An answer is required for each of the languages.) [L2]	In which of these languages were you fluent before age 7? (An answer is required for each of the languages.) [L3]	In which of these languages were you fluent before age 7? (An answer is required for each of the languages.) [L4]	In which of these languages were you fluent before age 7? (An answer is required for each of the languages.) [L5]	In which country/countries did you grow up?	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (An answer is required for each of the languages.) [mother tongue]	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (An answer is required for each of the languages.) [L2]	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (An answer is required for each of the languages.) [L3]
Französisch	Italienisch	Chinesisch	No	No	No	No	Deutschland	Yes	No	No
Französisch			No	No	No	No	Deutschland	Yes	Yes	No
Französisch	Schwedisch		No	No	No	No	Deutschland	Yes	No	No
Englisch	Spanisch	Arabisch	No	No	No	No	Deutschland	Yes	No	No
Spanisch			No	No	No	No	Deutschland	Yes	No	No
Französisch	Italienisch		No	No	No	No	Deutschland	Yes	No	No
Französisch	Spanisch		No	No	No	No	Deutschland	Yes	No	No
Spanisch	Italienisch		No	No	No	No	Deutschland	Yes	No	No
Spanisch	-	-	No	No	No	No	Deutschland	Yes	No	No

Which was the language of instruction for general classes during your primary and / or secondary education? (An answer is required for each of the languages.) [L4]	Which was the language of instruction for general classes during your primary and / or secondary education? (An answer is required for each of the languages.) [L5]	Module A2: Language acquisition (L2) At what age did you start learning L2?	How did you primarily learn L2?	How did you primarily learn L2? [Other]	For how many years have you been learning your L2?	Do you hold a language certificate for your L2 (e.g. TOEFL)?	If yes, which certificate?	If yes, at what age did you earn it?	Have you ever been to an L2 country for more than three months?
No	No	5	Formal education		15	No			No
No	No	8	Formal education		17	No			No
No	No	12	Formal education		11	No			No
No	No	9	Formal education		9	No			No
No	No	8	Formal education		15	Yes	SEFIC for Business English	20	Yes
No	No	9	Formal education		15	Yes	Englisch: LCCI - EFB Level 1	17	No
No	No	6	Formal education		16	No			No
No	No	10	Formal education		14	No			No

No	No	8	Formal education		15	No			Yes

If yes, how long was your longest stay (in months)?	If yes, when did you stay in that country and why?	With whom do you speak in your L2? (An answer is required for each of the options.) [Family]	With whom do you speak in your L2? (An answer is required for each of the options.) [Relatives]	With whom do you speak in your L2? (An answer is required for each of the options.) [Friends]	With whom do you speak in your L2? (An answer is required for each of the options.) [Colleagues/workplace]	How often do you speak in L2? Choose the option closest to your situation.	How many hours per week do you read texts in your mother tongue?
		Yes	Yes	Yes	No	From 1 to 15 hours a day	10
		No	No	Yes	Yes	From 1 to 15 hours a week	3
		No	No	Yes	Yes	From 1 to 15 hours a week	2
		No	No	No	No	From 1 to 15 hours a month	10
6	Praktikum im Rahmen des Bachelorstudiums	No	No	Yes	Yes	From 1 to 15 hours a day	4
		No	No	No	Yes	From 1 to 15 hours a month	15
		No	No	Yes	No	From 1 to 15 hours a day	5
		No	No	Yes	Yes	From 1 to 15 hours a day	3
6	September 2020 bis März 2021 für ein Praktikum	No	No	Yes	No	From 1 to 15 hours a week	6

How many hours per week do you read L2 texts?	How many hours per week do you listen to audio content in your mother tongue?	How many hours per week do you watch TV/movies/series in your mother tongue?	How many hours per week do you listen to audio content in L2?	How many hours per week do you watch TV/movies/series in L2?	Module A3 (Language acquisition): third language Do you speak a third language?	Which is your third language?	At what age did you start learning your third language?
35	0	0	0	10	Yes	Französisch	11
5	3	0	7	5	Yes	Französisch	16
3	0	10	4	3	Yes	Französisch	14
0	0	0	0	0	Yes	Französisch	12
4	2	4	1	3	Yes	Spanisch	14
5	5	12	3	7	Yes	Französisch	11
10	0	0	0	14	Yes	Französisch	11
3	0	0	0	12	Yes	Spanisch	15
10	4	2	10	15	Yes	Spanisch	17

How did you primarily learn this language?	How did you primarily learn this language? [Other]	For how many years have you been learning this language?	Do you hold a language certificate for your third language (e.g. TOEFL)?	If yes, which certificate was it?	If yes, at what age did you earn it?	Have you ever been to a country for more than three months where this language is spoken?	If yes, how long was your longest stay (in months)?
Formal education		9	No			No	
Formal education		6	No			No	
Formal education		9	Yes	DALF C1	21	Yes	4
In a country where this language is spoken		15	Yes	Master Französisch Lehramt	27	Yes	12
Formal education		5	No			No	
Formal education		13	Yes	DELF B1	15	Yes	5
Formal education		11	No			No	
Formal education		9	No			Yes	8
Formal education		5	No			No	

If yes, when did you stay in that country and why?	How do you rate your knowledge of your mother tongue? From 1 (very poor) to 100 (very good)	How do you rate your knowledge of your L2? From 1 (very poor) to 100 (very good)	How do you rate your active knowledge of your third language? From 1 (very poor) to 100 (very good)	How do you rate your passive knowledge of your third language? From 1 (very poor) to 100 (very good)
	100	90	80	100
	90	75	1	10
Auslandssemester in Lyon, ERASMUS, weil ich Frankreich liebe und dort leben wollte und weil es sehr gut zu meinem Studium gepasst hat.	80	70	70	70
2010 Au-Pair, 2019-20 Arbeit	90	70	75	85
	100	85	10	25
Sep 2014 - Jan 2015, Erasmus	95	80	70	75
	100	90	60	80
2016 für einen Europäischen Freiwilligendienst	80	50	50	65
	80	75	25	45

Module C1: Professional experience What is your current status?	If student, how is your course of studies called?	Do you already have a university degree?	If yes, please specify?	If student, in which year of study are you now?	What is your language combination?	Did you engage in any professional training before you started your degree program at the university?
Interpreting student	MAKD Englisch, Französisch	Yes	BA Sprache Kultur Translation	1	Englisch Französisch	No
Interpreting student	MAKD AB+Ü Englisch	Yes	MAKD AB+Ü Englisch	1	A-Sprache Deutsch, B-Sprache Englisch	No
Interpreting student	Master Konferenzdolmetschen A-Deutsch, B-Französisch, C-Englisch	Yes	Bachelor in Anwendungsorientierte Interkulturelle Sprachwissenschaft (Uni Augsburg)	1	A-Deutsch, B-Französisch, C-Englisch	No
Interpreting student	MAKD Deutsch Französisch Englisch	Yes	MA Ed	1	A: Deutsch B: Französisch C: Englisch	No

Interpreting student	Master Konferenzdolmetschen	Yes	Bachelor Fachübersetzen	1	Deutsch-Englisch	No
Interpreting student	Dolmetschen	Yes	Bachelor	1	DE - EN - FR	Yes
Interpreting student	Konferenzdolmetschen AB+Ü	Yes	Bachelor of Arts	1	A Deutsch B Englisch	No
Interpreting student	Dolmetschen mit Englisch und Übersetzen mit Englisch und Spanisch	Yes	Bachelor of Arts	1	Englisch und Spanisch	No
Interpreting student	MA Konferenzdolmetschen	Yes	BA Fachübersetzen (nach Abgabe der BA-Arbeit)	1	Englisch (ABÜ)	Yes

If yes, what type of training?	If yes, how long did the training last? (in months)	Did you gather any other interpreting experience? Please discribe.(For example interpreting classes, interpreting seminars, jobs etc.)	Have you ever interpreted for a friend or family member? (For example at the doctor or at the public administration office)	If so, where or in which context?	If so, how often or how regularly?	Are you interpreting beyond the context of your university studies?(For example as a freelancer alongside your studies)
			No			No
			No			No

		Im Rahmen der deutsch-französischen Partnerschaftsarbeit durfte ich beim alljährlichen Austausch zwischen meinem Heimatdorf und einem Dorf aus Frankreich konsekutiv Dolmetschen. Hauptsächlich Tischreden der beiden Bürgermeister. Das war alles ehrenamtlich.	No	Im Rahmen der deutsch-französischen Partnerschaftsarbeit durfte ich beim alljährlichen Austausch zwischen meinem Heimatdorf und einem Dorf aus Frankreich konsekutiv Dolmetschen. Hauptsächlich Tischreden der beiden Bürgermeister. Das war alles ehrenamtlich.	Das war vor Corona einmal pro Jahr für insgesamt drei Jahre.	No
		Nein	No			No
		2 Jahre Unterricht in Konsekutivdolmetschen und ca 4 Monate Simultanunterricht an der Würzburger Dolmetscherschule.	Yes	Für einen amerikanischen Austauschschüler, der ca. 1 Woche bei uns wohnte	Abgesehen davon, kaum.	No
Fremdsprachenkorrespondentin und danach staatlich geprüfte Übersetzerin	60	Propädeutikum während des BA, 2 Semester	Yes	Familientreffen	Bisher 3 Mal	No
		Dolmetschpropädeutikum Deutsch - Englisch	No			No
		Im Bachelorstudium habe ich schon ein Dolmetschpropädeutikum besucht	Yes	Während meines Europäischen Freiwilligendienstes habe ich bereits einige Erfahrungen im Community Interpreting gesammelt, als es darum ging, zwischen Freunden und Einheimischen sprachlich zu vermitteln.	zirka dreimal im Monat	No

Ausbildung zum staatlich geprüften Dolmetscher und Übersetzer	36	In meiner Ausbildung habe ich hauptsächlich Erfahrung im Gesprächs- und Konsekutivdolmetschen gesammelt. Vor der Pandemie hatte ich auch ein paar Monate Unterricht im Simultandolmetschen. Der Schwerpunkt der Ausbildung lag jedoch im Übersetzen.	No	-	-	No

Module C2: Interpreting competence How do you rate your ability to interpret from mother tongue into L2? From 1 (very poor) to 100 (very good)	How do you rate your ability to interpret from L2 into your mother tongue? From 1 (very poor) to 100 (very good)	How many hours per week do you engage in interpretation from your mother tongue to L2?	How many days per year do you engage in interpretation from your mother tongue to L2?	How many hours per week do you engage in interpretation from L2 to your mother tongue?	How many days per year do you engage in interpretation from L2 to your mother tongue?	What is the proportion of simultaneous vs. consecutive interpreting you engage in? [Simultaneous]	What is the proportion of simultaneous vs. consecutive interpreting you engage in? [Consecutive]	Are you member of the AIIC?
70	70	5	50	5	50	60	40	No
20	30	3	10	3	10	3	3	No
30	40	5	221	5	221	50	50	No
55	70	6	0	6	0	40	60	No
50	60	3	32	3	32	50	50	No
60	85	4	100	4	100	20	20	No
30	50	3	104	3	104	50	50	No
20	75	3	56	3	56	50	50	No
65	75	5	0	5	0	50	50	No

Module C3: Questions on the interpreting process Module C3.1: Preparation phase How often do you consider the interpreting brief before you start your preparation (in %)? [Not applicable to me.]	Module C3: Questions on the interpreting process Module C3.1: Preparation phase How often do you consider the interpreting brief before you start your preparation (in %)? [Other]	How often do you get customer-specific information (e.g., terminology, parallel texts)? (in %) [Not applicable to me.]	How often do you get customer-specific information (e.g., terminology, parallel texts)? (in %) [Other]	How often do you get any conference materials in advance? (in %) [Not applicable to me.]	How often do you get any conference materials in advance? (in %) [Other]	How often do you verify the speaker's first and/or provenance? (in %) [Not applicable to me.]	How often do you verify the speaker's first and/or provenance? (in %) [Other]	Do you prepare a glossary for the interpreting task? [In print]	Do you prepare a glossary for the interpreting task? [Digital: term database]
No	70	No	70	No	10	No	0	No	No
Yes		Yes		Yes		Yes		Yes	No
Yes		Yes		Yes		No	20	No	No
Yes		Yes		Yes		Yes		No	No
No	100	No	60	No	50	No	100	No	No
No	0	No	0	No	0	No	0	No	No
No	100	Yes		No	50	No	100	No	No
Yes		No	50	Yes		Yes		Yes	No

No	70	Yes		Yes		No	70	No	No

Do you prepare a glossary for the interpreting task? [Digital: Word list]	Do you prepare a glossary for the interpreting task? [no]	Module C3.2: Research strategies Do you consult external sources (websites, fora, databases, books) before the interpreting assignment?	If yes, ...	Do you consult own resources (e.g., glossaries, Translation Memories, term databases) before the interpreting assignment?	If yes, ...	During the research phase, what is your strategy if you know the meaning of a word, but you can't think of its most adequate equivalent in the target language? I …	Do you use translation memory systems?
No	Yes	No		No		use bilingual dictionary.	No
No	No	No		No		use bilingual dictionary.	No
Yes	No	Yes	Digital	No		use bilingual dictionary.	No
No	Yes	No		No		use bilingual dictionary.	No
Yes	No	Yes	Digital	Yes	Both	use monolingual dictionary.	No
Yes	No	Yes	Both	Yes	Both	infer the meaning from context.	No
Yes	No	Yes	Digital	No		use a search engine.	No

Yes	No	No		Yes	Both	use a search engine.	No
Yes	No	Yes	Digital	Yes	Digital	use a search engine.	No

Do you use terminology databases for your data management?	Do you use search engines like Google or ixquick?	What do you mainly look for when you use search engines?	Do you Google mainly in the source or target language?	Do you use interpreter-specific tools, e.g., Interplex or Interpretbank?	Which research aids do you use? [Monolingual Dictionaries]	Which research aids do you use? [Bilingual Dictionaries]	Which research aids do you use? [Synonym Dictionaries]	Which research aids do you use? [Encyclopedias or similar]	Which research aids do you use? [Search Engines]
No	No		target language	No	No	Yes	No	No	No
No	Yes	Background information	source language	Not applicable to me.	Yes	Yes	Yes	No	Yes
No	Yes	Background information	target language	Not applicable to me.	Yes	Yes	Yes	No	Yes
No	Yes	Background information	target language	No	Yes	Yes	Yes	No	Yes
No	Yes	Background information	both	Not applicable to me.	Yes	Yes	Yes	Yes	Yes
No	Yes	Background information	both	No	Yes	Yes	Yes	No	Yes
No	Yes	Words	both	No	No	No	No	Yes	Yes

No	Yes	Idiomatic expressions	both	Not applicable to me.	No	Yes	Yes	No	Yes
No	Yes	Idiomatic expressions	both	No	Yes	Yes	Yes	Yes	Yes

Which research aids do you use? [Terminology Databases]	Which research aids do you use? [Concordances Search Functions]	Which research aids do you use? [Not applicable to me]	Module C3.3: Simultaneous interpreting How important is professional equipment (e.g., interpreting booth)?from 0 (not important) - 100 (very important) [Not applicable to me.]	Module C3.3: Simultaneous interpreting How important is professional equipment (e.g., interpreting booth)?from 0 (not important) - 100 (very important) [Other]	How often do you work together with a partner? (in %) [Not applicable to me.]	How often do you work together with a partner? (in %) [Other]	If yes, does your interpreting partner help you while interpreting ?	If you work with a partner, after which interpreting duration do you usually switch over to your partner (in minutes)?	What is your maximum simultaneous interpreting capacity per day? (in hours)
No	No	No	No	100	No	0	No	0	2
No	No	No	No	100	Yes		N/A		1,5
No	No	No	No	50	Yes		N/A		0,5
No	No	No	Yes		Yes		N/A		0
Yes	Yes	No	No	80	Yes		N/A		1,5
No	No	No	No	100	Yes		N/A		4,5

No	No	No	No	80	Yes		N/A		1,5
No	No	No	No	100	Yes		N/A		1,5
No	No	No	No	70	Yes		N/A		2

What is your estimated decalage? (in seconds) [Not applicable to me]	What is your estimated decalage? (in seconds) [Other]	How important is visual contact to the speaker during simultaneous interpreting? from 0 (not important) - 100 (very important) [Not applicable to me.]	How important is visual contact to the speaker during simultaneous interpreting? from 0 (not important) - 100 (very important) [Other]	Do you pay attention to any other visual information distributed by the speaker (e.g., presentations, slides, handouts)?	What strategy do you use when you don’t understand the speaker acoustically?	What strategy do you use when you don’t understand the meaning of a term or sentence? [Ask my partner]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Omission of the passage]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Guessing the content]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Invent new contents]
No	3	No	60	Yes	Omission of the passage	No	Yes	Yes	No
No	4	No	50	Yes	Not applicable to me	No	Yes	No	No

No	2	No	10	Yes	Omission of the passage	No	Yes	Yes	No
Yes		Yes		Not applicable to me	Not applicable to me	No	No	No	No
Yes		No	100	Yes	Not applicable to me	No	No	No	No
No	4	No	70	Yes	Omission of the passage	No	Yes	Yes	Yes
No	3	No	50	Yes	Omission of the passage	No	No	No	No
No	7	No	30	Yes	Invent new contents	No	No	Yes	Yes
No	4	No	60	Yes	Omission of the passage	No	No	Yes	No

What strategy do you use when you don’t understand the meaning of a term or sentence? [Interrupt the speaker]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Look up in resource]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Stalling]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Not applicable to me]	What strategy do you use when you can't think of a term in the target language?	How do you deal with spontaneous changes in the manuscript, if you get one?	Do you correct obvious mistakes of the speaker in the target language?	Do you use any resources in the booth? [Online resources]
No	No	Yes	No	Invent new contents	Interpret the changes	Yes	No
No	No	Yes	No	Omission of the passage	Interpret the changes	Yes	No

No	Yes	Yes	No	Guessing the content	Not applicable to me	Yes	No
No	No	No	Yes	Not applicable to me	Not applicable to me	Not applicable to me	No
No	No	Yes	No	Not applicable to me	Interpret the changes	Yes	No
No	No	No	No	Stalling	Not applicable to me	Not applicable to me	No
No	No	Yes	No	Invent new contents	Not applicable to me	Not applicable to me	No
No	No	No	No	Guessing the content	Not applicable to me	Not applicable to me	No
No	No	No	No	Guessing the content	Interpret the changes	Yes	No

Do you use any resources in the booth? [Term database]	Do you use any resources in the booth? [TM]	Do you use any resources in the booth? [Google]	Do you use any resources in the booth? [Print glossaries]	Do you use any resources in the booth? [Digital word lists]	Do you use any resources in the booth? [Interpreter-specific tools, e.g. Interplex or Interpretbank]	Do you use any resources in the booth? [No]	Do you use any resources in the booth? [Not applicable to me]	Module C3.4: Consecutive interpreting How often do you use note taking strategies? (in %) [Not applicable to me]	Module C3.4: Consecutive interpreting How often do you use note taking strategies? (in %) [Other]
No	No	No	No	No	No	Yes	No	No	100
No	No	No	No	No	No	No	Yes	No	10
No	No	No	No	No	No	No	Yes	No	95
No	No	No	No	No	No	No	Yes	Yes	
No	No	No	No	No	No	Yes	No	No	90
No	No	No	Yes	No	No	No	No	No	90
No	No	No	No	No	No	Yes	No	No	100

No	No	No	Yes	No	No	No	No	No	30
No	No	Yes	Yes	Yes	No	No	No	No	90

Do you have your own notation system?	Do you omit articles, exclamations, unimportant prepositions, adjectives and adverbs while taking notes?	Do you separate your notes by ideas?	What is the speaker’s maximum speaking time before you have to start the interpreting task? (in minutes) [Not applicable to me]	What is the speaker’s maximum speaking time before you have to start the interpreting task? (in minutes) [Other]	What is your maximum interpreting capacity per day? (in hours) [Not applicable to me]	What is your maximum interpreting capacity per day? (in hours) [Other]	What strategy do you use when you don’t understand the speaker acoustically?	What strategy do you use when you don’t understand the meaning of a term or sentence?
Yes	Yes	Yes	No	10	No	3	Omission of the passage	Omission of the passage

No	Not applicable to me	No	No	5	No	1,5	Omission of the passage	Omission of the passage
Yes	Yes	Yes	No	4	Yes		Ask the speaker to repeat	Ask the speaker to explain
N/A			Yes		Yes		Not applicable to me	Not applicable to me
Yes	No	Not applicable to me	No	6	No	1,5	Ask the speaker to repeat	Not applicable to me
Yes	Yes	Yes	No	5	No	3	Omission of the passage	Guessing the content
Yes	Yes	Yes	No	4	No	1,5	Guessing the content	Guessing the content
Yes	Yes	Yes	No	5	No	1	Guessing the content	Invent new contents
No	Yes	No	No	8	No	2	Omission of the passage	Guessing the content

What strategy do you use when you can't think of a term in the target language?	Do you pay attention to any other visual information distributed by the speaker (e.g., presentations, slides, handouts)?	Do you correct obvious mistakes of the speaker during consecutive interpreting?	How do you deal with inappropriate or offensive statements?	Module C3.5: Quality management How do you assess the quality of your interpreting?	How often do you get feedback from your client? (in %) [Not applicable to me]	How often do you get feedback from your client? (in %) [Other]	If so, how do you deal with this feedback?
I use synonyms	Yes	Yes	Literal translation	Feedback by colleague	Yes		
Omission of the passage	Yes	Yes	Manipulated translation	Recordings	Yes		
Guessing the content	Yes	Yes	Not applicable to me	Recordings	Yes		
Not applicable to me	Not applicable to me	Not applicable to me	Not applicable to me	Recordings	Yes		

Not applicable to me	Yes	Yes	Not applicable to me	Recordings	Yes		
I use synonyms	Yes	Not applicable to me	Not applicable to me	Recordings	Yes		
Invent new contents	Yes	Not applicable to me	Manipulated translation	Recordings	No	60	I make notes for my next job
Guessing the content	Yes	Not applicable to me	Manipulated translation	Feedback by colleague	Yes		
Guessing the content	Yes	Yes	Manipulated translation	Recordings	Yes		

Post-Test TICQ Experimental group

Module C2: Interpreting competence How do you rate your ability to interpret from mother tongue into L2? From 1 (very poor) to 100 (very good)	How do you rate your ability to interpret from L2 into your mother tongue? From 1 (very poor) to 100 (very good)	How many hours per week do you engage in interpretation from your mother tongue to L2?	How many days per year do you engage in interpretation from your mother tongue to L2?	How many hours per week do you engage in interpretation from L2 to your mother tongue?	How many days per year do you engage in interpretation from L2 to your mother tongue?	What is the proportion of simultaneous vs. consecutive interpreting you engage in? [Simultaneous]	What is the proportion of simultaneous vs. consecutive interpreting you engage in? [Consecutive]	Are you member of the AIIC?
90	90	4	4	4	4	50	50	No
1	85	6	12	4	12	60	40	No
75	85	4	40	3	40	50	50	No
70	80	8	50	8	50	30	20	No
75	80	5	150	5	150	60	40	No
60	70	7	50	7	50	4	1	No

76	82	15	15	0	0	50	50	No
20	30	3	24	3	24	50	50	No

Module C3: Questions on the interpreting process Module C3.1: Preparation phase How often do you consider the interpreting brief before you start your preparation (in %)? [Not applicable to me.]	Module C3: Questions on the interpreting process Module C3.1: Preparation phase How often do you consider the interpreting brief before you start your preparation (in %)? [Other]	How often do you get customer-specific information (e.g., terminology, parallel texts)? (in %) [Not applicable to me.]	How often do you get customer-specific information (e.g., terminology, parallel texts)? (in %) [Other]	How often do you get any conference materials in advance? (in %) [Not applicable to me.]	How often do you get any conference materials in advance? (in %) [Other]	How often do you verify the speaker's first and/or provenance? (in %) [Not applicable to me.]	How often do you verify the speaker's first and/or provenance? (in %) [Other]	Do you prepare a glossary for the interpreting task? [In print]	Do you prepare a glossary for the interpreting task? [Digital: term database]
No	100	No	80	Yes		Yes		No	No
No	100	Yes		No	100	Yes		No	No
No	10	No	30	No	20	Yes		No	No
No	100	Yes		Yes		Yes		No	No
No	100	Yes		Yes		No	100	Yes	No

No	100	No	60	No	30	No	100	Yes	No
No	78	No	80	No	90	No	40	Yes	No
Yes		No	30	Yes		Yes		Yes	No

Do you prepare a glossary for the interpreting task? [Digital: Word list]	Do you prepare a glossary for the interpreting task? [no]	Module C3.2: Research strategies Do you consult external sources (websites, fora, databases, books) before the interpreting assignment?	If yes, ...	Do you consult own resources (e.g., glossaries, Translation Memories, term databases) before the interpreting assignment?	If yes, ...	During the research phase, what is your strategy if you know the meaning of a word, but you can't think of its most adequate equivalent in the target language? I …	Do you use translation memory systems?	Do you use terminology databases for your data management?
Yes	No	Yes	Digital	Yes	Digital	use bilingual dictionary.	No	No
Yes	No	Yes	Both	Yes	Both	use a search engine.	No	No
No	Yes	Yes	Digital	Yes	Digital	use bilingual dictionary.	No	No
Yes	No	Yes	Both	Yes	Digital	infer the meaning from context.	No	No
Yes	No	Yes	Both	Yes	Both	use a search engine.	No	No

Yes	No	Yes	Both	Yes	Both	use a search engine.	No	No
No	No	Yes	Digital	No		use a search engine.	No	No
No	No	No		Yes	Print	use bilingual dictionary.	No	No

Do you use search engines like Google or ixquick?	What do you mainly look for when you use search engines?	Do you Google mainly in the source or target language?	Do you use interpreter-specific tools, e.g., Interplex or Interpretbank?	Which research aids do you use? [Monolingual Dictionaries]	Which research aids do you use? [Bilingual Dictionaries]	Which research aids do you use? [Synonym Dictionaries]	Which research aids do you use? [Encyclopedias or similar]	Which research aids do you use? [Search Engines]
Yes	Words	target language	No	No	Yes	No	No	Yes
Yes	Background information	both	No	Yes	Yes	Yes	Yes	Yes
Yes	Background information	both	No	Yes	Yes	Yes	Yes	Yes
Yes	Background information	target language	No	Yes	Yes	Yes	Yes	Yes
Yes	Background information	both	Not applicable to me.	No	Yes	No	Yes	Yes
Yes	Idiomatic expressions	target language	No	Yes	Yes	Yes	Yes	Yes
Yes	Phrases	both	No	Yes	Yes	No	Yes	Yes

Yes	Idiomatic expressions	both	No	No	Yes	No	No	Yes

Which research aids do you use? [Terminology Databases]	Which research aids do you use? [Concordances Search Functions]	Which research aids do you use? [Not applicable to me]	Module C3.3: Simultaneous interpreting How important is professional equipment (e.g., interpreting booth)?from 0 (not important) - 100 (very important) [Not applicable to me.]	Module C3.3: Simultaneous interpreting How important is professional equipment (e.g., interpreting booth)?from 0 (not important) - 100 (very important) [Other]	How often do you work together with a partner? (in %) [Not applicable to me.]	How often do you work together with a partner? (in %) [Other]	If yes, does your interpreting partner help you while interpreting?	If you work with a partner, after which interpreting duration do you usually switch over to your partner (in minutes)?	What is your maximum simultaneous interpreting capacity per day? (in hours)
No	No	No	No	90	No	5	Yes	15	4
No	Yes	No	No	90	No	100	Yes	30	2
No	No	No	No	80	No	5	Yes	20	4
Yes	Yes	No	No	100	Yes		N/A		3
No	No	No	No	100	Yes		N/A		1
No	No	No	No	100	No	10	Yes	20	2

No	No	No	Yes		Yes		N/A		3
No	No	No	Yes		Yes		N/A		0,5

What is your estimated decalage? (in seconds) [Not applicable to me]	What is your estimated decalage? (in seconds) [Other]	How important is visual contact to the speaker during simultaneous interpreting? from 0 (not important) - 100 (very important) [Not applicable to me.]	How important is visual contact to the speaker during simultaneous interpreting? from 0 (not important) - 100 (very important) [Other]	Do you pay attention to any other visual information distributed by the speaker (e.g., presentations, slides, handouts)?	What strategy do you use when you don’t understand the speaker acoustically?	What strategy do you use when you don’t understand the meaning of a term or sentence? [Ask my partner]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Omission of the passage]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Guessing the content]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Invent new contents]
No	4	No	50	Yes	Omission of the passage	Yes	No	Yes	No

No	5	No	60	Yes	Omission of the passage	Yes	Yes	Yes	No
No	3	No	10	Yes	Ask my partner	Yes	Yes	No	No
No	3	No	100	Yes	Not applicable to me	No	Yes	No	No
No	3	No	90	Yes	Omission of the passage	Yes	Yes	No	No
No	2	No	80	Yes	Omission of the passage	No	No	No	No
No	9	No	70	Yes	Not applicable to me	Yes	Yes	Yes	No
No	10	Yes		Yes	Guessing the content	No	No	Yes	No

What strategy do you use when you don’t understand the meaning of a term or sentence? [Interrupt the speaker]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Look up in resource]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Stalling]	What strategy do you use when you don’t understand the meaning of a term or sentence? [Not applicable to me]	What strategy do you use when you can't think of a term in the target language?	How do you deal with spontaneous changes in the manuscript, if you get one?	Do you correct obvious mistakes of the speaker in the target language?	Do you use any resources in the booth? [Online resources]	Do you use any resources in the booth? [Term database]
No	Yes	Yes	No	Look up in resource	Interpret the changes	Yes	Yes	No

No	No	Yes	No	Stalling	Interpret the changes	Yes	Yes	No
No	Yes	Yes	No	Stalling	Interpret the changes	Not applicable to me	Yes	No
No	Yes	Yes	No	Look up in resource	Interpret the changes	Yes	Yes	No
No	No	Yes	No	Stalling	Interpret the changes	Yes	Yes	No
No	Yes	Yes	No	Look up in resource	Interpret the changes	Yes	Yes	No
No	No	Yes	No	Stalling	Interpret the changes	Yes	No	No
No	No	No	No	Guessing the content	Not applicable to me	Not applicable to me	No	No

Do you use any resources in the booth? [TM]	Do you use any resources in the booth? [Google]	Do you use any resources in the booth? [Print glossaries]	Do you use any resources in the booth? [Digital word lists]	Do you use any resources in the booth? [Interpreter-specific tools, e.g. Interplex or Interpretbank]	Do you use any resources in the booth? [No]	Do you use any resources in the booth? [Not applicable to me]	Module C3.4: Consecutive interpreting How often do you use note taking strategies? (in %) [Not applicable to me]	Module C3.4: Consecutive interpreting How often do you use note taking strategies? (in %) [Other]
No	Yes	No	Yes	No	No	No	No	100
No	Yes	No	Yes	No	No	No	No	95
No	Yes	No	Yes	No	No	No	No	5
No	Yes	Yes	Yes	No	No	No	No	100
No	Yes	No	Yes	No	No	No	No	100

No	Yes	Yes	Yes	No	No	No	No	100
No	No	Yes	Yes	No	No	No	No	100
No	No	No	No	No	No	Yes	No	30

Do you have your own notation system?	Do you omit articles, exclamations, unimportant prepositions, adjectives and adverbs while taking notes?	Do you separate your notes by ideas?	What is the speaker’s maximum speaking time before you have to start the interpreting task? (in minutes) [Not applicable to me]	What is the speaker’s maximum speaking time before you have to start the interpreting task? (in minutes) [Other]	What is your maximum interpreting capacity per day? (in hours) [Not applicable to me]	What is your maximum interpreting capacity per day? (in hours) [Other]	What strategy do you use when you don’t understand the speaker acoustically?	What strategy do you use when you don’t understand the meaning of a term or sentence?	What strategy do you use when you can't think of a term in the target language?
Yes	Yes	Yes	No	15	No	1	Omission of the passage	Guessing the content	I use synonyms

Yes	Yes	Yes	No	7	No	2	Ask the speaker to repeat	Omission of the passage	I use synonyms
No	Yes	No	No	10	No	4	Ask the speaker to repeat	Ask the speaker to explain	I use synonyms
No	Yes	Yes	No	10	Yes		Ask the speaker to repeat	Omission of the passage	I use synonyms
Yes	Yes	Yes	No	8	No	1	Ask the speaker to repeat	Ask the speaker to explain	I use synonyms
Yes	Yes	Yes	No	2	No	2	Guessing the content	Omission of the passage	I use synonyms
Yes	Yes	No	No	5	No	3	Guessing the content	Guessing the content	I use synonyms
No	No	No	No	5	No	0,25	Guessing the content	Guessing the content	Guessing the content

Do you pay attention to any other visual information distributed by the speaker (e.g., presentations, slides, handouts)?	Do you correct obvious mistakes of the speaker during consecutive interpreting?	How do you deal with inappropriate or offensive statements?	Module C3.5: Quality management How do you assess the quality of your interpreting?	How often do you get feedback from your client? (in %) [Not applicable to me]
Yes	Yes	Literal translation	Recordings	Yes
Yes	Yes	Manipulated translation	Feedback by colleague	Yes

Yes	Not applicable to me	Not applicable to me	Recordings	Yes
Yes	Yes	Not applicable to me	Feedback by client	Yes
Yes	Yes	Manipulated translation	Feedback by colleague	Yes
Yes	Yes	Manipulated translation	Feedback by colleague	Yes
Yes	Yes	Not applicable to me	Recordings	Yes
Yes	No	Not applicable to me	I don´t	Yes

Baseline TICQ Control Group[6]

Please enter the participant ID.	Please enter the name of your study: EEGHa2021	Module A1: Personal data Age:	Gender:	What is your highest degree?	What is your highest degree? [Other]	Which languages do you speak? [Mother tongue (L1)]	Which languages do you speak? [Language 2 (L2)]	Which languages do you speak? [Language 3 (L3)]	Which languages do you speak? [Language 4 (L4)]	Which languages do you speak? [Language 5 (L5)]	In which of these languages were you fluent before age 7? (An answer is required for

---

[6] Only the aspects that are relevant to the between group-analysis at baseline are included

											each of the languages.) [L2]
P12	EEGHa2021	23	Female	Bachelor		Deutsch	Englisch	Französisch	Italienisch		No
P13	EEGHa2021	23	Female	Bachelor		Deutsch	Englisch	Französisch	Italienisch		No
P14	EEGHa2021	35	Female	Bachelor		Französisch	Englisch	Deutsch	Spanisch		No
P16	EEGHa2021	21	Female	Bachelor		Deutsch	Englisch	Französisch			No
P17	EEGHa2021	22	Female	Bachelor		polish	Englisch	german	french		No
P18	EEGHa2021	22	Female	Bachelor		Deutsch	Arabisch	Englisch	Spanisch		Yes
P20	EEGHa2021	25	Female	Bachelor		Greek	Albanian	English	German		Yes
P21	EEGHa2021	23	Female	Secondary school		Deutsch	Englisch	Französisch			No
P22	EEGHa2021	23	Female	Bachelor		Chinesisch	Englisch	Deutsch			No
P23	EEGHa2021	23	Female	Bachelor		Polish	German	English	French		Yes
P24	EEGHa2021	24	Male	Bachelor		German	English	Spanish	Italian		No
P25	EEGHa2021	23	Male	Bachelor		Russian	English	German	Spanish		No

In which of these languages were you fluent before age 7? (An answer is required for each of the	In which of these languages were you fluent before age 7? (An answer is required for each of the	In which of these languages were you fluent before age 7? (An answer is required for each of the	In which country/countries did you grow up?	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (A	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (A	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (A	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (A	Which was the language of instruction for general classes (e.g., maths, biology, etc.) during your primary and / or secondary education? (A	Module A2: Language acquisition (L2) A t what age did you start learning L2?

languages.) [L3]	languages.) [L4]	languages.) [L5]		n answer is required for each of the languages.) [mother tongue]	n answer is required for each of the languages.) [L2]	n answer is required for each of the languages.) [L3]	n answer is required for each of the languages.) [L4]	n answer is required for each of the languages.) [L5]	
No	No	No	Deutschland	Yes	No	No	No	No	7
No	No	No	Deutschland	Yes	No	No	No	No	6
No	No	No	Kamerun	Yes	No	No	No	No	6
No	No	No	Deutschland	Yes	No	No	No	No	10
No	No	No	Poland	Yes	No	No	No	No	6
Yes	No	No	Deutschland-Ägypten	No	Yes	Yes	No	No	0
No	No	No	Greece	Yes	No	No	No	No	0
No	No	No	Deutschland	Yes	No	No	No	No	7
No	No	No	China	Yes	No	No	No	No	16
No	No	No	Germany	No	Yes	No	No	No	3
No	No	No	Germany	Yes	No	No	No	No	9
No	No	No	Russian Federation	Yes	No	No	No	No	8

How did you primarily learn L2?	How did you primarily learn L2? [Other]	For how many years have you been learning your L2?	Do you hold a language certificate for your L2 (e.g. TOEFL)?	If yes, which certificate?	If yes, at what age did you earn it?	Have you ever been to an L2 country for more than three months?	If yes, how long was your longest stay (in months)?
Formal education		16	No			No	
Formal education		13	Yes	TOEFL ITP	17	Yes	5

In an L2 speaking country		18	No			Yes	6
Formal education		11	No			Yes	6
Formal education		16	No			No	
In an L2 speaking country		10	No			Yes	120
Parents		25	No			No	
Formal education		16	No			Yes	5
Formal education		12	No			No	
Formal education		20	No			Yes	23
Formal education		16	Yes	Cambridge Certificate	16	No	
Formal education		15	No			No	

If yes, when did you stay in that country and why?	With whom do you speak in your L2? (An answer is required for each of the	With whom do you speak in your L2? (An answer is required for each of the	With whom do you speak in your L2? (An answer is required for each of the	With whom do you speak in your L2? (An answer is required for each of the options.) [Colleagues/workplace]	How often do you speak in L2? Choose the option closest to your situation.	How many hours per week do you read texts in your mother tongue?	How many hours per week do you read L2 texts?

	options.) [Family]	options.) [Relatives]	options.) [Friends]				
	No	No	No	Yes	From 1 to 15 hours a week	2	4
2019 als Au-Pair und 2021 im Auslandssemester	No	No	Yes	Yes	From 1 to 15 hours a week	5	3
Es ist mein Heimatland	Yes	Yes	Yes	No	From 1 to 15 hours a month	4	3
Erasmus+, August 2022 bis Januar 2023	No	No	Yes	Yes	From 1 to 15 hours a day	20	10
	No	No	Yes	No	From 1 to 15 hours a week	7	6
Umzug der Eltern	Yes	Yes	Yes	No	From 1 to 15 hours a day	7	4
	Yes	Yes	No	No	From 1 to 15 hours a month	8	1
Erasmus; 09/2022-01/2023	No	No	Yes	Yes	From 1 to 15 hours a week	21	21
	No	No	No	Yes	From 1 to 15 hours a month	3	1
i was born in Germany	Yes	No	Yes	Yes	From 1 to 15 hours a day	5	30
	No	No	Yes	No	From 1 to 15 hours a month	3	2
	No	No	Yes	Yes	From 1 to 15 hours a month	18	9

How many hours per week do you listen to audio content in your	How many hours per week do you watch TV/movies/series in your mother tongue?	How many hours per week do you listen to audio content in L2?	How many hours per week do you watch TV/movies/series in L2?	Module A3 (Language acquisition): third language D o you speak a third language?	Which is your third language?	At what age did you start learning your third language?	How did you primarily learn this language?

mother tongue?							
0	3	0	5	Yes	Französisch	12	Formal education
4	4	8	3	Yes	Französisch	10	Formal education
2	4	2	4	Yes	Deutsch	13	Formal education
50	10	40	20	Yes	Französisch	11	Formal education
10	1	10	2	Yes	german	13	Formal education
1	12	0	5	Yes	Englisch	8	Formal education
2	1	0	0	Yes	English	7	Formal education
5	10	5	15	Yes	Französisch	12	Formal education
1	10	0	3	Yes	Deutsch	16	Formal education
0	5	2	0	Yes	English	7	Formal education
2	0	4	8	Yes	Spanish	16	In a country where this language is spoken
8	16	8	12	Yes	German	14	Formal education

For how many years have you been learning	Do you hold a language certificate for your third language	If yes, which certificate was it?	If yes, at what age did you earn it?	Have you ever been to a country for more than three	If yes, how long was your longest stay (in months)?	If yes, when did you stay in that country and why?

this language?	(e.g. TOEFL)?			months where this language is spoken?		
11	No			No		
9	No			No		
13	Yes	TestDaf	23	Yes	150	zum Zweck des Studiums
10	Yes	diplôme du baccalauréat général	17	No		
9	Yes	DSD B1	15	No		
10	No			No		
9	Yes	Cambridge	16	No		
11	No			No		
7	Yes	TestDAF	22	Yes	23	07.2021-06.2023
15	No			No		
9	No			Yes	12	Mexico, I did a gap year there working as a volunteer
9	Yes	Goethe Zertifikat C2	22	No		

How do you rate your knowledge of your mother tongue?	How do you rate your knowledge of your L2? From 1 (very	How do you rate your active knowledge of your third	How do you rate your passive knowledge of your third	Module C1: Professional experience What is your current status?	If student, how is your course of studies called?	Do you already have a university degree?

From 1 (very poor) to 100 (very good)	poor) to 100 (very good)	language? From 1 (very poor) to 100 (very good)	language? From 1 (very poor) to 100 (very good)			
95	80	40	60	Bilingual without any interpreting training		N/A
98	90	75	85	student	Translation	Yes
100	80	90	100	student	Translation, Master of Arts	Yes
90	60	40	50	student	Master Translation	Yes
85	70	50	65	student	Translations-, Sprach- und Kulturwissenschaft	Yes
85	80	65	65	student	Fachübersetzen	No
100	70	90	95	Bilingual without any interpreting training		N/A
95	80	40	60	Bilingual without any interpreting training		N/A
95	60	75	75	student	Übersetzen	Yes
85	90	70	80	Bilingual without any interpreting training		N/A
90	85	80	85	student	MA Translation	Yes
88	80	80	76	Bilingual without any interpreting training		N/A

If yes, please specify?	If student, in which year of study	What is your language combination?	Did you engage in any professional training before you

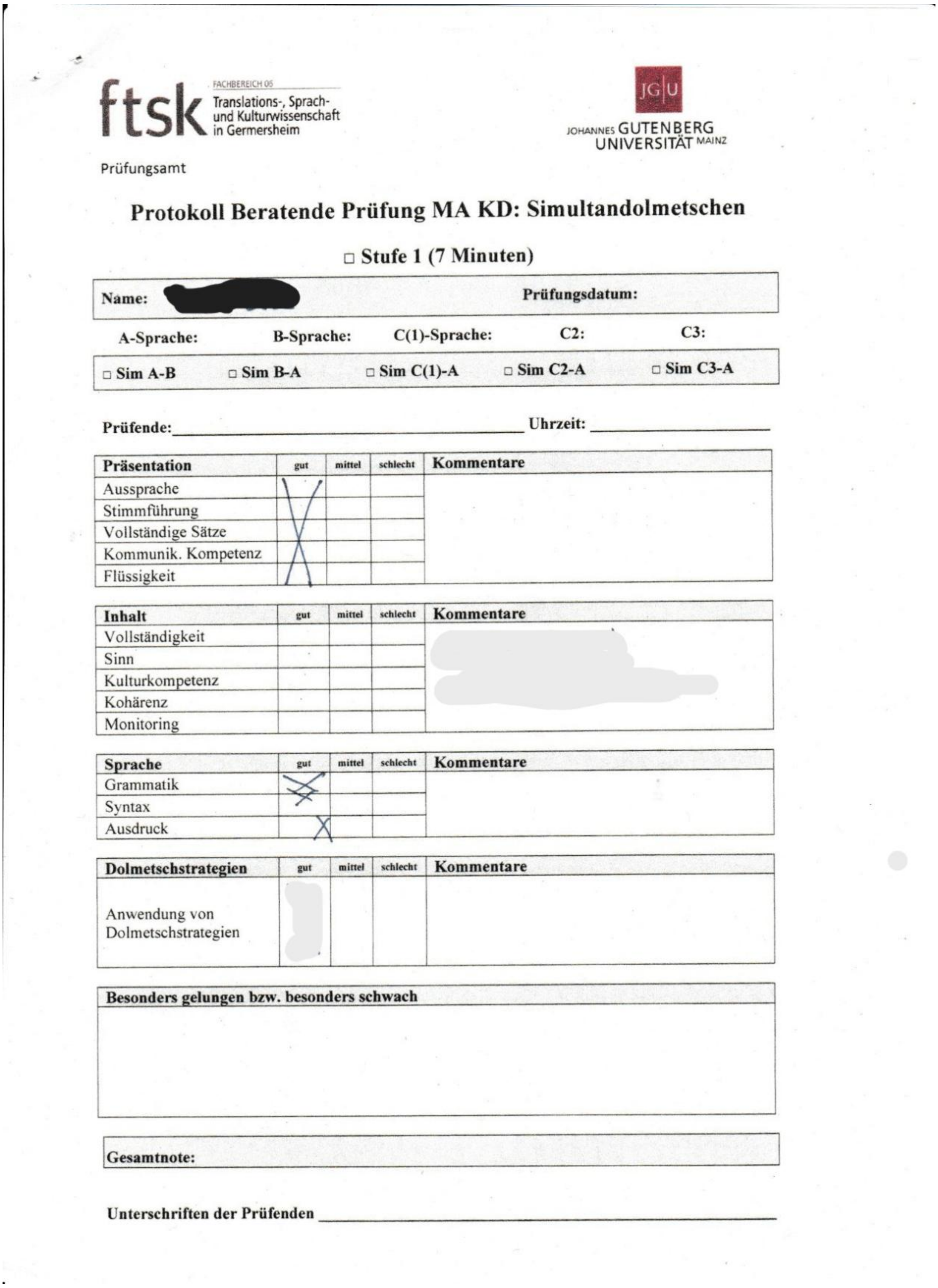
	are you now?		started your degree program at the university?
			N/A
B.A. Sprach-, Kultur- und Translationswissenschaften (EN, FR, IT)	8	Deutsch (G), Englisch, Französisch	No
Translation, Bachelor of Arts	1	Grundsprache Französisch, F1 Deutsch, F2 Englisch	No
Bachelor	7	Grundsprache Deutsch, F1 Französisch, F2 Englisch	No
BA Applied linguistics	1	polish, german, english	No
	1	Deutsch, Arabisch, Spanisch	No
			N/A
			N/A
Uni Bonn Volkswirtschaftslehre	1	Chinesisch und Deutsch	No
			N/A
BA Sprache, Kultur, Translation	5	Spanisch and Englisch	No
			N/A

Did you gather any other interpreting experience? Please discribe.(For example interpreting classes, interpreting seminars, jobs etc.)	Have you ever interpreted for a friend	If so, where or in which context?	If so, how often or how regularly?	Are you interpreting beyond the context of your

	or family member?			university studies?
no				
no				
no				
no				
no	yes	to enable conversations between my mum and my friends	every couple of months	no
no				
no				
no				
no	Ja	In China. Ich habe meinem Vater geholfen, mit Ausländer zu kommunizieren.	nur einmal	Nein
interpreting classes in my bachelor degree and unprofessional interpreting experience for friends and family	yes	doctors appointments, school gatherings, friends vistings, administrative affairs	in my childhood/teenage years more often	no
I took one interpreting class from German to Spanish, but I didn't like it as interpreting made me feel very overwhelmed and stressed.	For my boyfriend.	When my boyfriend came to Germany, I had to be the interpreter between him and my parents because he doesn't speak German or English.	Just for a couple of days.	No
I take interpreting classes as an elective discipline (Wahlpflichtmodul) at the moment but I have never taken such classes before. Also, I don't have any interpreting work experience.	No			No

Module C2: Interpreting competence How do you rate your ability to	How do you rate your ability to interpret from L2 into your

interpret from mother tongue into L2? From 1 (very poor) to 100 (very good)	mother tongue? From 1 (very poor) to 100 (very good)
1	1
10	25
80	100
1	1
70	75
60	70
80	90
20	50
0	0
60	50
15	25
50	50



ftsk FACHBEREICH 06 Translations-, Sprach- und Kulturwissenschaft in Germersheim

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Prüfungsamt

Protokoll Beratende Prüfung MA KD: Simultandolmetschen

☐ Stufe 1 (7 Minuten)

Name: [redacted] **Prüfungsdatum:**

A-Sprache: **B-Sprache:** **C(1)-Sprache:** **C2:** **C3:**

☐ **Sim A-B** ☐ **Sim B-A** ☐ **Sim C(1)-A** ☐ **Sim C2-A** ☐ **Sim C3-A**

Prüfende:________________ **Uhrzeit:** ________

Präsentation	gut	mittel	schlecht	**Kommentare**
Aussprache	X			
Stimmführung				
Vollständige Sätze				
Kommunik. Kompetenz				
Flüssigkeit				

Inhalt	gut	mittel	schlecht	**Kommentare**
Vollständigkeit				[redacted]
Sinn				
Kulturkompetenz				
Kohärenz				
Monitoring				

Sprache	gut	mittel	schlecht	**Kommentare**
Grammatik	X			
Syntax	X			
Ausdruck	X			

Dolmetschstrategien	gut	mittel	schlecht	**Kommentare**
Anwendung von Dolmetschstrategien	[redacted]			

Besonders gelungen bzw. besonders schwach

Gesamtnote:

Unterschriften der Prüfenden ________________

Eigenständigkeitserklärung

Hiermit erkläre ich, dass diese Dissertation selbstständig, ohne fremde Hilfe und mit keinen anderen als den darin angegebenen Hilfsmitteln angefertigt wurde, dass die wörtlichen oder dem Inhalt nach aus fremden Arbeiten entnommenen Stellen, Zeichnungen, Skizzen, bildliche Darstellungen und dergleichen als solche genau kenntlich gemacht sind und dass von der Ordnung zur Sicherung guter wissenschaftlicher Praxis in Forschung und Lehre und zum Verfahren zum Umgang mit wissenschaftlichem Fehlverhalten Kenntnis genommen wurde. Zudem erkläre ich, dass keine Hilfe von kommerziellen Promotionsberatern in Anspruch genommen wurde.

Ort, Datum: Germersheim, 07.05.2025 Unterschrift: