From oscillations to language: behavioural and electroencephalographic studies on cross-language interactions

Von Grebmer Zu Wolfsthurn, S.

CHAPTER 1

General introduction

1.1 Background

How does the brain manage multiple languages? What are the underlying behavioural and neural mechanisms of the brain trying to make sense of the world from input that is not in the native language (L1), or to produce output in a non-native language? Anecdotally, when *late language learners* are asked about their strategy to name a particular object in a non-native language, they will often state that they rely on the similarities between their L1 and the non-native language in question. In this thesis, we defined late language learners as those individuals who have acquired a language in addition to their L1 after the age of fourteen (S. Rossi, Gugler, Friederici & Hahne, 2006). We placed a special focus on late language learners due to an noticeable lack of research combining behavioural and neurolinguistic methods when examining this particular population. Moreover, we broadly defined individuals who are able to actively engage and communicate in more than one language as *multilinguals* (Kroll, Dussias, Bice & Perrotti, 2015). For those individuals to be considered multilingual, proficiency levels can vary in terms of their comprehension and speaking proficiency but also within their L1 and any additional languages.

Current evidence overwhelmingly suggests that multilingualism has profound functional and structural consequences for the mind and brain and that it impacts both language-specific mechanisms and more domain-general cognitive mechanisms (Abutalebi, Cappa & Perani, 2001; Bialystok, Craik & Luk, 2012; Kroll et al., 2015; Stein et al., 2012; Wong, Yin & O'Brien, 2016). For example, previous research suggested a differential use of domain-general cognitive control networks for multilinguals compared to monolinguals (Abutalebi et al., 2013; Bialystok et al., 2012; D. W. Green & Abutalebi, 2013). Cognitive control refers to the ability to adjust to a preferred outcome, for example suppressing irrelevant information in a particular context (Abutalebi & Green, 2016; Declerck et al., 2021; D. W. Green, 1998). Multilinguals were consistently found to outperform monolinguals on tasks which required the inhibition of irrelevant information, the switching of attention and management of working memory load (Hilchey & Klein, 2011). Critically, in this thesis we move away from the traditional comparison between multilinguals and monolinguals. Instead, we focus on multilingualism as a gateway to the underlying processing mechanisms of the brain with respect to non-native language comprehension and production. In light of this, we aim to add novel evidence to the current literature (Abutalebi et al., 2001; Berthele, 2021a; Kroll & Bialystok, 2013; Kroll et al., 2015; Puig-Mayenco, González Alonso & Rothman, 2020).

A central theme of this thesis is cross-linguistic influence, or *CLI*, which encompasses the bidirectional influence on processing mechanisms of the L1 and non-native language at the cognitive and neural level (Lago, Mosca & Garcia, 2021; Nozari & Pinet, 2020). Weinreich (1953) was one of the first to describe CLI as a key feature in the context of multilingual language processing (Kroll et al., 2015). CLI speaks directly to the anecdotal accounts mentioned before, whereby learners describe that they form "connections" between their L1 and a non-native language. Given its potentially critical role in non-native language processing, CLI has since featured in several studies which made fundamental contributions to the characterisation of the multilingual language architecture (Clahsen & Felser, 2006a; Hartsuiker, Pickering & Veltkamp, 2004; Jarvis, 2011; Odlin, 1989; B. D. Schwartz & Sprouse, 1996). CLI has its roots in the parallel activation of the languages within the multilingual brain, also in cases where only one language is required in a specific context (Blumenfeld & Marian, 2007; Costa & Pickering, 2019; Guo & Peng, 2006; Kroll et al., 2015; Marian & Spivey, 2003a; Mishra & Singh, 2016). Yet, CLI and its impact on non-native comprehension and production has been scarcely researched in late language learners with lower proficiency levels in comparison to highly proficient bilinguals (Bordag & Pechmann, 2007; Bürki & Laganaro, 2014; Costa, Kovacic, Franck & Caramazza, 2003; Lemhöfer, Spalek & Schriefers, 2008; Morales et al., 2016). Therefore, in this thesis we thoroughly examined two CLI effects in late language learners in more detail, namely the *gender congruency effect*, and the *cognate facilitation effect*.

The gender congruency effect refers to the notion of more efficient processing of gender congruent nouns (i.e., nouns with similar gender categories across languages) vs. incongruent nouns (i.e., nouns with dissimilar gender categories across languages) as a result of the interaction of the grammatical gender systems (Morales et al., 2016; Paolieri, Demestre, Guasch, Bajo & Ferré, 2020). Note that we exclusively focused on grammatical gender as opposed to semantic gender, which is based on the biological features of the noun (Corbett, 1991). Our main language of interest in this thesis was Spanish, which is characterised by a two-way gender system. The definite determiner "la" marks the feminine gender value, and "el" marks masculine gender value, e.g., la_F $mesa_F$ [the table] and el_M $perro_M$ [the dog], respectively. Contrastingly, the cognate facilitation effect refers to the processing advantage of cognates (i.e., words with considerable orthographic and phonological overlap) compared to non-cognates, for example *el volcán - the volcano* vs. *el perro - the dog* in Spanish and English (Amengual, 2012; Bosma, Blom, Hoekstra & Versloot, 2019; Costa, Caramazza & Sebastián-Gallés, 2000; Dijkstra, Miwa, Brummelhuis, Sappelli & Baayen, 2010; Lemhöfer

& Dijkstra, 2004; Strijkers, Costa & Thierry, 2010). As mentioned above, these two CLI effects are robustly reported for highly proficient multilinguals (Bordag & Pechmann, 2007; Lemhöfer et al., 2008; Morales et al., 2016), but have been less extensively studied in late language learners. Therefore, by focusing on late language learners, we obtained critical insights from CLI into multilingual language processing in a relatively understudied population. Among those insights were the quantification of CLI in non-native comprehension and production, the sensitivity of the multilingual brain to the corresponding linguistic features of gender congruency and cognate status, the neural correlates of CLI in late language learners, the characterisation of individual non-native production stages as well as the locus of target language selection. These issues are systematically featured in this thesis across a number of studies, as discussed in more detail in the following sections.

The second central theme of this thesis is whether and how *language similarity* between the L1 and the non-native language affects non-native comprehension and production mechanisms. Defining and measuring language similarity has been subject to debate in the literature (Ringbom & Jarvis, 2009; Van der Slik, 2010). Previous work attempted to quantify language similarity by means of genetic linguistic distance (Cavalli-Sforza, Menozzi & Piazza, 1994), or cognate linguistic distance (McMahon & McMahon, 2005; Van der Slik, 2010). While we acknowledge that this debate around how to exactly measure the similarity between two languages is far from resolved, in this thesis we defined language similarity as the structural morphosyntactic, orthographic and phonological word form overlap across languages (Foote, 2009; Rothman & Cabrelli Amaro, 2010). Note also that throughout this thesis, we will use the term language similarity interchangeably with the terms linguistic similarity and typological similarity. Following this definition, we tested three different multilingual populations with varying degrees of language similarity: native German late learners of Spanish and native Dutch late learners of Spanish, representing the linguistically more dissimilar language pairs; and native Italian late learners of Spanish, representing the linguistically more similar language pair.

Beyond the aforementioned anecdotal accounts, research suggested that learners routinely embed the non-native language in their L1 context during language processing (Ringbom & Jarvis, 2009). Yet, language similarity was only featured in a few studies on non-native language comprehension and production (Dijkstra et al., 2010; Foucart & Frenck-Mestre, 2011; Sabourin, Stowe & De Haan, 2006; Rothman & Cabrelli Amaro, 2010; Zawiszewski, Gutiérrez, Fernández & Laka, 2011; Zawiszewski & Laka, 2020). Evidence from these studies tentatively suggested a processing advantage for speakers of linguistically similar languages, e.g., Italian-Spanish, compared to speakers of linguistically less similar languages, e.g., German-Spanish (Sabourin et al., 2006; Zawiszewski & Laka, 2020). The debate around the role of language similarity and its impact on multilingual processing is critical because it taps directly into the co-existence of two or more languages in a multilingual brain and the subsequent functional and cognitive architecture of each of these languages (Abutalebi, 2008; Clahsen & Felser, 2006b; Tolentino & Tokowicz, 2011). In light of the apparent lack of research, characterising the exact impact of language similarity on language processing mechanisms is a crucial issue. Therefore, in this thesis we quantified the role language similarity plays in multilingual language processing to gain insights into how language similarity guides non-native processing. More specifically, we tested the influence of high or low language similarity on non-native comprehension and production in Italian-Spanish speakers and German-Spanish speakers across several studies. The critical question here was whether or not speakers of highly similar languages, e.g., Italian-Spanish, had an inherent processing advantage compared to speakers of less similar languages, e.g., German-Spanish.

Additionally, we explored whether language similarity affected higher cognitive functioning such as cognitive control in Italian-Spanish speakers and in Dutch-Spanish speakers. Here, a fundamental issue is whether speakers of linguistically similar languages, e.g., Italian-Spanish, develop enhanced cognitive control skills over time compared to speakers of linguistically less similar languages, e.g., Dutch-Spanish (Stocco, Yamasaki, Natalenko & Prat, 2014;

Yamasaki, Stocco & Prat, 2018). Taken together, we investigated the role of language similarity both in the context of non-native comprehension and production, but also in the context of more general cognitive functions such as cognitive control.

1.2 Non-native processing through the lens of electroencephalography

In this thesis, we combined behavioural measures of performance accuracy, response times, naming accuracy and naming latencies with electrophysiological measures, i.e., electroencephalography (EEG). EEG is a non-invasive method to measure event-related brain potentials (ERPs). These ERPs are time-locked to particular events, e.g., the appearance of a stimulus during a task (Woodman, 2010). More specifically, ERPs reflect the systematic oscillatory changes of brain potentials over time. These systematic changes are termed ERP components and can be linked to particular linguistic and cognitive phenomena, for example syntactic or semantic processing, or cognitive control (Friederici, Steinhauer & Frisch, 1999; Steinhauer, White & Drury, 2009; Swaab, Ledoux, Camblin & Boudewyn, 2011). ERPs are typically measured via scalp electrodes in the form of voltage amplitudes. Due to their excellent temporal resolution, ERPs can provide detailed and nuanced insights into the sub-mechanisms of language processing as well as domain-general cognitive mechanism in real time (Bürki & Laganaro, 2014; Christoffels, Firk & Schiller, 2007; Foucart & Frenck-Mestre, 2011; Hahne & Friederici, 2001; E. M. Moreno, Rodríguez-Fornells & Laine, 2008; S. Rossi et al., 2006; Steinhauer et al., 2009; Tokowicz & MacWhinney, 2005; Valente, Bürki & Laganaro, 2014; Woodman, 2010). Therefore, EEG and ERPs are fundamental tools in examining non-native language comprehension and production mechanisms. In this thesis, we focused mainly on the exploration of two ERP components: the P600 component in (non-)native comprehension in Chapters 2, 4, 5 and 8; and the P300 component in non-native production in Chapters 3, 4 and 6. However, we also briefly touched on the left anterior negativity (LAN) component, and the N400 component in Chapters 2, 3 and 8; both of which were commonly linked to language comprehension mechanisms (Barber & Carreiras, 2005; Kutas & Federmeier, 2011; Martín-Loeches, Muñoz, Casado, Melcón & Fernández-Frías, 2005; Swaab et al., 2011).

The P600 component is typically elicited in the context of syntactic violation paradigms (Hagoort, Brown & Groothusen, 1993). It has a maximal peak over centro-parietal regions around 600 ms post-stimulus onset and a component latency between 500 ms to 900 ms post-stimulus onset (Steinhauer et al., 2009; Swaab et al., 2011). Importantly, P600 component amplitudes vary with respect to whether a linguistic structure contains a syntactic violation: evidence suggests larger voltage amplitudes for syntactically incorrect compared to syntactically correct constructions, yielding the so-called *P600 effect* (Swaab et al., 2011). The P600 effect was robustly reported across a range of language combinations and different types of syntactic violations, for example determiner-noun agreement violations or adjective-noun agreement violations (Barber & Carreiras, 2005; Friederici et al., 1999; Gunter, Friederici & Schriefers, 2000; Hagoort & Brown, 2000; Molinaro, Vespignani & Job, 2008; Osterhout & Holcomb, 1992; Steinhauer et al., 2009; Swaab et al., 2011). The consensus is that the P600 component is a neural correlate of successful syntactic integration, re-analysis and repair (Alemán Bañón, Fiorentino & Gabriele, 2012; Barber & Carreiras, 2005; Gouvea, Phillips, Kazanina & Poeppel, 2010). Relevant to this thesis, the P600 effect was less robustly reported for late language learners with lower proficiency levels (Hahne, 2001; Hahne & Friederici, 2001; Weber-Fox & Neville, 1996), but see S. Rossi et al. (2006). This tentatively suggested that late language learners showed decreased sensitivity to syntactic violations. Therefore, in this thesis we used the P600 component to first, shine light onto the underlying neural mechanisms of (non-)native comprehension. Second, we examined how factors such as language similarity or CLI could impact P600 component voltage amplitudes. Third, we studied whether late language learners demonstrated sensitivity to syntactic violations. Finally, we examined whether P600 component amplitudes differed as a function of language similarity. These issues were closely investigated in Chapters 2, 4, 5 and 8.

The second ERP component we focused on in this thesis was the P300 component. This component is frequently reported in centro-parietal regions with a peak around 300 ms after stimulus onset and was more generally associated with cognitive control, working memory load and inhibitory processes (Barker & Bialystok, 2019; Barry et al., 2020; González Alonso et al., 2020; Polich, 2007). The so-called *P300 effect* refers to higher voltage amplitudes elicited in the more cognitively demanding condition, for example the incongruent condition in a classical Flanker task (Eriksen & Eriksen, 1974). This particular task exploits the mismatch between a target stimulus and the surrounding stimuli, e.g., a target arrow in the centre of an array where the remaining arrows point in a different direction to the target arrow (Bosma & Pablos, 2020). More recently, the P300 component was also reported in designs combining both cognitive control and language processing components, see for example Bosma and Pablos (2020) for a Flanker task with intermittent code-switching. Yet, to this date not much research has exploited the potential of the P300 component to study key features such as CLI in non-native language production. Even fewer studies have tackled the time course of non-native production with respect to CLI (Bürki & Laganaro, 2014; Valente et al., 2014). Therefore, in this thesis we used the P300 component to explore how CLI modulated the time course of non-native production. We used the Levelt-Roelofs-Meyer (LRM) model (Levelt, Roelofs & Meyer, 1999) as our theoretical framework. This model describes word production in terms of several individual production stages in the L1, each connected to a specific duration in time (Indefrey & Levelt, 2004; Indefrey, 2011). In this thesis, we utilised the gender congruency effect and the cognate facilitation effect to examine two different theoretical accounts of the mitigation of CLI and the selection of the target language. One account suggested that CLI was resolved prior to the so-called lexical retrieval stage (Gollan, Montoya, Fennema-Notestine & Morris, 2005; Hermans, Bongaerts, De Bot & Schreuder, 1998), whereas a second account suggested that CLI was only resolved after lexical retrieval (Christoffels et al., 2007; Colomé, 2001; Rodriguez-Fornells et al., 2005). Therefore, a novel aspect of this thesis was the examination of the P300 component as a potential index of mechanisms critical to successful non-native language production, namely the mitigation of CLI between the L1 and the non-native language and target language selection. Moreover, we explored whether P300 component amplitudes differed across various populations with a higher or lower language similarity. These issues are examined in detail in Chapters 3, 4 and 6 of this thesis. Taken together, combined with adequate statistical tools, EEG and ERPs form powerful pillars in the quest of disentangling the complex processing architecture within the multilingual brain.

1.3 Open Science and statistics in neurolinguistics

In recent years, there have been growing concerns about the robustness, reproducibility and interpretation of published research findings within the scientific community (Sönning & Werner, 2021). These concerns marked the beginnings of the so-called *replication crisis* in fields such as psycholinguistics or neurolinguistics, but also in linguistics, psychology and neuroscience (Ioannidis, 2005; Shrout & Rodgers, 2018; Sönning & Werner, 2021; Tackett, Brandes, King & Markon, 2019). Sönning and Werner (2021) outlined four problematic aspects of published research in an attempt to capture these concerns: the high rates of false-positive results connected to the (mis)use of statistical techniques, the opaqueness in terms of the methodological details and the analysis procedures, the inaccessibility of the original data, and finally, the perceived insignificance of replication studies and their subsequent classification as "non-original" research. Several studies have made efforts to define the underlying causes of these concerns, which include the incorrect interpretation of statistical outcome parameters, sub-optimal research designs and low statistical power (Ioannidis, 2005; Munafò et al., 2017). On the other hand, research has also focused on defining mitigatory strategies of the replication crisis, such as making data and research materials freely accessible and elaborating on the selection of statistical analysis procedures (Sönning & Werner, 2021). This movement is now widely known as Open Science, see Mirowski (2018) and Vicente-Saez and Martinez-Fuentes (2018) for a detailed discussion.

In the current thesis, we continued these efforts and placed a strong focus on the following: first, we increased the transparency regarding the linguistic background of our multilingual participants, the research method and the analysis procedures. More specifically, in addition to our strong theoretical focus, we provided detailed descriptions of our participants, research design, stimuli and materials. Moreover, we meticulously described and motivated our statistical analysis approaches and reported the detailed model specifications and parameters. Second, we made the data and analysis scripts openly available to the scientific community by creating public repositories on the Open Science Framework, OSF (Foster & Deardorff, 2017). This was done in an effort to encourage replication studies and to be transparent about the data and the analysis details. Third, we went beyond traditional statistical analysis approaches, especially for our highly complex EEG data. This meant that for some of our data, we expanded on more conventional analysis approaches and applied relatively novel, but more suitable statistical techniques. This was done to challenge current analysis practises and to propose concrete alternative approaches. To provide tangible examples, we used more advanced statistical methods to analyse both our behavioural data and our EEG data on a single-trial basis: (generalised) linear mixed effects models (LMMs), see Frömer, Maier and Abdel Rahman (2018); and generalised additive mixed models (GAMMs), see De Cat, Klepousniotou and Baayen (2015), Meulman, Wieling, Sprenger, Stowe and Schmid (2015) and Tremblay and Newman (2015).

An important characteristic of LMMs is the ability to effectively estimate random variance, and to manage missing data and unbalanced datasets (Baayen, Davidson & Bates, 2008; Fröber, Stürmer, Frömer & Dreisbach, 2017). As a result, LMMs yield robust estimates while minimising false-positives and maintaining high statistical power (Barr, 2013; Matuschek, Kliegl, Vasishth, Baayen & Bates, 2017). At the centre of this approach is a linear mixed model, which consists of a fixed effects structure and a random effects structure. The fixed effects structure typically contains the intercept or grand mean across all predictors, the linear predictors and the corresponding interaction effects, if applicable, as well as any potential covariates in line with the experimental design, research question and hypotheses. In contrast, the random effects structure contains terms to capture by-participant and by-item variance, and to estimate the effect of the main predictor variables as a function of the individual subject and item. LMMs are suitable for behavioural analyses, but are also a viable alternative to the more conventional ANOVA-approach for EEG data analysis (Frömer et al., 2018). LMMs for behavioural analyses are featured in all chapters of this thesis, and LMMs for EEG data analyses are included in Chapters 2, 3, 4 and 6 of this thesis.

On the other hand, the generalised additive mixed models, or *GAMMs*, represent an extension of LMMs: in addition to linear predictors, GAMMs also include non-linear predictors which do not assume a linear relationship between the predictor and the outcome variable (Meulman et al., 2015; Tremblay & Newman, 2015). These non-linear terms are fundamentally a collection of so-called "basis functions", which for instance capture model voltage amplitudes over time (De Cat et al., 2015; Meulman et al., 2015). This is a particularly critical feature of GAMMs, making them optimal candidates for analysing the complex and multi-dimensional oscillatory trends of EEG data. From the researcher's perspective, this approach avoids the a priori specification of a time window of interest for a given effect. Instead, it provides a precise estimate for the time window of this effect. Critically, GAMMs also include random intercepts and random slopes to capture random variance, by-participant and by-item effects, which is similar to LMMs. Therefore, GAMMs are a powerful approach to analysing EEG data compared to more conventional methods, in particular when examining populations that are less frequently featured in the current literature. In this thesis, we used GAMMs to flexibly model our EEG data in Chapters 5 and 8.

1.4 The current thesis

As previously outlined, the critical issues of this thesis were the following: first, to examine and quantify cross-linguistic influence (CLI) in both non-native comprehension and production. Second, to characterise the corresponding neural correlates of CLI in the multilingual brain. Third, to explore the sensitivity of late language learners to syntactic irregularities. Fourth, to examine the individual non-native production stages and the selection of the target language in light of CLI. Fifth, to investigate the overarching impact of language similarity between the native language and the non-native language on non-native comprehension and production. Finally, to study the impact of language similarity on cognitive control. Subsequently, our core research questions were concerned with the characterisation of the impact of both CLI and language similarity on the underlying processing mechanisms in late language learners, both from a behavioural and from a neural perspective. These issues were explored across a series of experiments, which are outlined separately in the corresponding chapters. We used a variety of tasks and analysis techniques within and across several participant groups, namely German-Spanish speakers, Italian-Spanish speakers and Dutch-Spanish speakers, to tap directly into each one of our critical issues and the corresponding research questions.

In **Chapter 2** of this thesis, we examined the effects of CLI on non-native comprehension in a syntactic violation paradigm in native German late learners of Spanish. More specifically, we studied how two CLI effects, namely the gender congruency effect and the cognate facilitation effect, modulated participants' ability to correctly identify gender agreement violations in determiner-noun pairs, for example [el pato] vs. [*la pato] *the duck*. In this, we also sought to characterise the neural correlates of syntactic violation processing in the form of the P600 component. This was because previous research on late language learners was scarce and tended to suggest a decreased sensitivity of late learners to syntactic violations. Moreover, we also tested whether and how non-native vocabulary size was related to cognitive control, which is a central feature of managing CLI between two languages. This study therefore contributed to current research first, with insights into the cumulative impact of CLI in late language learners and the potential interaction effect between linguistic features of gender congruency and cognate status on non-native comprehension; next, with a characterisation of late language learners' sensitivity to syntactic violations; and finally, with a description of CLI in non-native comprehension from a neural perspective.

In **Chapter 3**, we reported a study on non-native production conducted with the same German-Spanish speakers from Chapter 2. Non-native production is characterised by CLI as well as the necessity to select a target language over a non-target language. Yet, little research has been done on the impact of CLI on the time course of non-native production. Therefore, we used two CLI effects, namely the gender congruency effect and the cognate facilitation effect, to model the temporal unfolding of the non-native production stages and their corresponding neural signatures. In other words, we examined how CLI impacted the timing of the individual production stage during an overt picture-naming task. Next, we also probed the locus of target language selection in late language learners, with a particular focus on the modulation of the P300 component during this process. The critical contributions of this particular study were a nuanced description of CLI effects during the individual non-native production stages, as well as how these effects impacted the time course of non-native production. Moreover, we tapped not only into the mitigation of CLI, but also into the locus of target language selection and the related neural correlates. Therefore, our findings were directly relevant for the characterisation of the time course of non-native production as well as for the debate around the selection of the target language in late language learners.

Chapter 4 used the previous two chapters as its starting point and examined the same issues surrounding CLI in non-native comprehension and production in a linguistically more similar language pair, namely native Italian late learners of Spanish. In other words, we used an identical theoretical framework and experimental design to determine whether the results from Chapters 2 and 3 would apply to speakers of highly similar languages (Italian and Spanish). Using a syntactic violation paradigm to study non-native comprehension, we investigated how gender congruency and cognate status as representatives of CLI influenced non-native comprehension mechanisms. Next, we studied whether the Italian-Spanish speakers also showed a sensitivity to syntactic violations, as reflected in the P600 effect. Finally, we asked whether the P600 component was modulated by CLI. To examine non-native production, we used a picture-naming task to probe impact of CLI on the time course of non-native production and the locus of target language selection. We were particularly interested in a potential modulation of the P300 component by CLI. Taken together, our basic question in this chapter was whether the findings from the previous two chapters would also prevail with respect to a different multilingual population.

In **Chapter 5** we systematically examined an effect of language similarity on CLI in a syntactic violation paradigm in late language learners with respect to non-native comprehension. More specifically, we brought together the German-Spanish and Italian-Spanish speakers studied in Chapters 2 and 4 to ask the fundamental questions of whether speakers of highly similar languages had a processing advantage over speakers of less similar languages, and whether this notion was traceable both in terms of behaviour and neural correlates. Subsequently, we examined first, whether there were distinct neural correlates (i.e., P600 effects) for syntactic violation processing in speakers of linguistically similar vs. linguistically less similar languages; and second, whether CLI effects at the level of gender and cognates varied as a function of language similarity. Crucially, this study allowed us to investigate language similarity both from the perspective of differential neural signatures across groups, but also in terms of CLI effects and what it meant to be a speaker of highly similar languages at the cognitive level. Therefore, this work had direct implications for characterising the relevance of language similarity in non-native comprehension. Moreover, it contributed novel evidence to the small pool of literature tackling this particular issue.

Chapter 6 complemented the previous chapter as it included the direct comparison of the German-Spanish and Italian-Spanish speakers from Chapters 3 and 4 in terms of non-native production. Previous work on language similarity effects in production is scarce, in particular in combination with electrophysiological measures. Therefore, the central aim of this study was to investigate a potential language similarity effect on CLI in speakers of linguistically similar languages (Italian-Spanish) and in speakers of linguistically less similar languages (German-Spanish) at the behavioural and neural level. In this, we also investigated how gender congruency and cognate status impacted overt picture-naming in late language learners. Similar to Chapters 3 and 4, we particularly focused on the modulation of the P300 component as a function of both language similarity and CLI. The findings from this study were relevant for the in-depth examination of the P300 component as an index for CLI mitigation, and spoke directly to the importance of language similarity in non-native production.

Chapter 7 went above and beyond the role of language similarity in non-native language processing. Instead, the study reported in this chapter examined the impact of language similarity on higher cognitive functions. We described a study exploring the effect of language similarity on inhibitory control performance using a spatial Stroop task (Hilbert, Nakagawa, Bindl & Bühner, 2014). This task is commonly employed to quantify cognitive control in the form of inhibitory control (Stroop, 1935). Expanding on the notion of measurable effects of multilingualism on cognitive control, we asked whether speaking highly similar languages, e.g., Italian and Spanish, had direct consequences for inhibitory control performance compared to speaking less similar languages, e.g., Dutch and Spanish. Therefore, the fundamental goal of this study was to explore whether or not speakers of similar languages (Italian-Spanish) yielded superior inhibitory control skills compared to speakers of less similar languages (Dutch-Spanish). This study not only had implications for the relative impact of language similarity on higher cognitive functioning, but also tapped into the broader notion of language-specific vs. domain-general consequences of multilingualism on the multilingual mind and brain.

The foundations for the study reported in **Chapter 8** were developed after feedback provided by an anonymous reviewer in Chapter 2. This particular reviewer highlighted the controversial nature of the ERP correlates of syntactic violation processing in native speakers of Spanish in the context of isolated gender agreement violations. Intrigued by the disparity of results with respect to the elicitation of ERP components during gender agreement violation processing, we therefore closely examined the corresponding neural correlates in a population of native Spanish speakers. Notably, this is the only study in this thesis focusing on L1 language comprehension. Here, we also placed a particular emphasis on using more advanced statistical methods to model the oscillatory tendencies of the EEG signal over time. Subsequently, this study contributed novel evidence to the debate around which could be the primary ERP component linked to gender agreement processing in Spanish. The study further examined to which extent some of the evidence we observed was related to the amount of context given to readers to identify a gender violation. Finally, the study tapped into the broader question of whether there are functionally and neuronally distinct ERP signatures for language-specific phenomena across different (native) languages.

Finally, **Chapter 9** served the purpose of integrating the findings from each of the studies reported in this thesis. Therefore, in this chapter we synthesised our research findings, their theoretical implications and future directions to form a more general picture of the question we asked at the very beginning of this thesis: how *does* the brain manage multiple languages?